

**НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ТЕЛЕКОМУНІКАЦІЙ І ГЛОБАЛЬНОГО
ІНФОРМАЦІЙНОГО ПРОСТОРУ**

Кваліфікаційна наукова праця
на правах рукопису

ПРОСЯНКІНА-ЖАРОВА ТЕТЯНА ІВАНІВНА

УДК 004.5:004.6:004.89:007.51

**ДИСЕРТАЦІЯ
ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ДЛЯ СИСТЕМ ПІДТРИМКИ
ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ**

05.13.06 – Інформаційні технології

подається на здобуття наукового ступеня
доктора технічних наук

Дисертація містить результати власних досліджень. Використання результатів і текстів інших авторів мають посилання на відповідне джерело

Науковий консультант:
член-кореспондент НАН України,
д.т.н., професор **Олександр ТРОФИМЧУК**

Київ – 2025

АНОТАЦІЯ

Присянкіна-Жарова Т. І. - Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – інформаційні технології. Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, 2025.

Дисертаційну роботу присвячено вирішенню актуальної науково-прикладної проблеми - розроблення інформаційних технологій, призначених для використання у системах підтримки прийняття рішень управління регіональним розвитком органів місцевого самоврядування та державного управління.

У дисертації розроблено нові інформаційні технології, які дозволяють обробляти великі обсяги різномірної структурованої та неструктурованої, часто недостовірної, обмеженої та неповної інформації, приймати коректні рішення щодо розвитку соціально-еколого-економічних систем – регіонів та громад в умовах невизначеності та ризиків різних типів.

У роботі представлено інформаційні технології побудови прогнозів розвитку соціально-еколого-економічних систем на основі системної методології адаптивного моделювання і комплексного застосування статистично-ймовірнісного підходу та методів інтелектуального аналізу даних. Розроблені інформаційні технології ґрунтуються на принципах багатомодельного підходу, інтеграції структурованої та неструктурованої інформації, системному використанні методів інтелектуального аналізу даних, моделювання, прогнозування та прийняття рішень. Науковими результатами досліджень є методології створення інформаційної технології виконання системних досліджень складних соціально-економічних систем, розроблення методологічних засад побудови на її основі інтелектуальної системи підтримки прийняття рішень для підвищення якості прийняття управлінських рішень щодо їх розвитку в умовах невизначеності.

В роботі розроблено системну методологію попередньої обробки (підготовки) даних, яка ґрунтується на множині методів відбору найбільш значимих чинників, заповнення пропусків даних і нормування спостережень, обробки екстремальних значень та структурування даних.

У дисертаційному дослідженні удосконалено системну методологію побудови моделей процесів, що характерні для різних рівнів складних соціально-економічних систем на основі адаптивного підходу до моделювання з комбінованим використанням сценарного аналізу, регресійних та ймовірнісно-статистичних моделей у формі мереж Байєса, що дозволяє урахувати структурно-параметричні невизначеності і забезпечує адекватний опис причинно-наслідкових зв'язків та можливих варіантів розвитку подій під впливом різних груп чинників;

У роботі представлено удосконалену системну методологію проєктування і реалізації систем підтримки прийняття рішень, що ґрунтується на принципах мережецентричного, багатомодельного та багатокритеріального підходів і засновану на системному використанні методів аналізу даних, адаптивного моделювання, прогнозування і прийняття рішень; що дозволяє органам державного управління та місцевого самоврядування вирішувати задачі регіонального розвитку, ефективно реагувати на загрози в умовах наявності невизначеностей та ризиків різних типів.

Запропоновано використання різних типів інформаційних технологій прогнозування, в основу яких покладено поетапне розкриття невизначеностей різної природи, уточнення результатів моделювання на кожному етапі дослідження. В основу запропонованої технології покладено багатомодельний підхід, який ґрунтується на використанні множини різнотипних моделей: регресійний аналіз, байєсівське моделювання і прогнозування, технологія на основі застосування методу аналізу подібності часових рядів, аналізу структурованих та неструктурованих даних. При цьому кожний тип моделей призначений для виконання поставленої конкретної задачі. Згідно з принципами багатомодельного підходу розроблено метод синтезу інформаційних технологій,

застосований для розв'язування задач прогнозування розвитку соціально-економічних систем, який ґрунтується на інтеграції різнотипної інформації й заснований на системному використанні методів аналізу даних, моделювання, методів прогнозування.

Запропонована інформаційна технологія інтерооперабельною. Програмне забезпечення реалізоване мовами програмування SAS/Base та SAS/IML та Python і має гнучку адаптивну архітектуру. Методи, що були розроблені та реалізовані в роботі, призначені насамперед для автоматизації аналізу та прогнозування процесів регіонального розвитку.

Запропоновані методи і моделі доведені до рівня практичної реалізації у вигляді інформаційної технології що дозволяє аналізувати та прогнозувати регіональний розвиток в умовах невизначеності.

Отримані теоретичні та практичні результати дисертаційної роботи використовуються в науково-практичній діяльності підприємств, установ, органів державного управління та місцевого самоврядування.

Ключові слова: інформаційна технологія, система підтримки прийняття рішень, мережецентричний підхід, невизначеність, регіональний розвиток, прогнозування, математичні моделі, штучний інтелект.

ABSTRACT

Dissertation for the degree of Doctor of Technical Sciences in the specialty 05.13.06 "Information Technologies". - Institute of Telecommunications and Global Information Space, National Academy of Sciences of Ukraine, Kyiv, 2025.

The dissertation is dedicated to solving a topical scientific and applied problem - the development of information technologies intended for use in decision-making support systems for regional development management of local governments and state administration.

The dissertation develops new information technologies that allow processing large volumes of heterogeneous structured and unstructured, often unreliable, limited and incomplete information, and making correct decisions regarding the development of socio-ecological and economic systems - regions and communities in conditions of uncertainty and risks of various types.

The paper presents information technologies for building forecasts of the development of socio-ecological and economic systems based on the systematic methodology of adaptive modeling and the integrated application of the statistical-probabilistic approach and methods of intelligent data analysis. The developed information technologies are based on the principles of a multi-model approach, integration of structured and unstructured information, systematic use of methods of intelligent data analysis, modeling, forecasting and decision-making. The scientific results of the research are methodologies for creating information technology for performing systematic research of complex socio-economic systems, developing methodological principles for building an intelligent decision-making support system based on it to improve the quality of management decision-making regarding their development in conditions of uncertainty.

The paper has developed a systematic methodology for pre-processing (preparation) of data, which is based on a set of methods for selecting the most significant factors, filling data gaps and normalizing observations, processing extreme values and structuring data.

The dissertation research improves the systemic methodology for building models of processes that are characteristic of different levels of complex socio-economic systems based on an adaptive approach to modeling with the combined use of scenario analysis, regression and probabilistic-statistical models in the form of Bayesian networks, which allows taking into account structural-parametric uncertainties and provides an adequate description of cause-and-effect relationships and possible scenarios under the influence of various groups of factors;

The work presents an improved systemic methodology for designing and implementing decision-making support systems, which is based on the principles of network-centric, multi-model and multi-criteria approaches and is based on the systematic use of data analysis methods, adaptive modeling, forecasting and decision-making; which allows state and local government bodies to solve regional development problems and effectively respond to threats in the presence of uncertainties and risks of various types.

The use of different types of forecasting information technologies is proposed, which are based on the gradual disclosure of uncertainties of different nature, refinement of modeling results at each stage of the study. The proposed technology is based on a multi-model approach, which is based on the use of a set of different types of models: regression analysis, Bayesian modeling and forecasting, technology based on the application of the method of time series similarity analysis, analysis of structured and unstructured data. At the same time, each type of model is designed to perform a specific task. In accordance with the principles of the multi-model approach, a method of information technology synthesis has been developed, applied to solving problems of forecasting the development of socio-economic systems, which is based on the integration of different types of information and is based on the systematic use of data analysis, modeling, and forecasting methods. The proposed information technology is interoperable. The software is implemented in the SAS/Base and SAS/IML and Python programming languages and has a flexible adaptive architecture. The methods that were developed and implemented in the work are intended primarily for the automation of the analysis and forecasting of regional development processes.

The proposed methods and models have been brought to the level of practical implementation in the form of information technology that allows analyzing and forecasting regional development under conditions of uncertainty.

The obtained theoretical and practical results of the dissertation work are used in the scientific and practical activities of enterprises, institutions, state administration bodies and local governments.

Keywords: information technology, decision support system, network-centric approach, uncertainty, regional development, forecasting, mathematical models, artificial intelligence.

Список опублікованих праць за темою дисертації

Наукові праці, в яких опубліковані основні наукові результати дисертації

Монографії

1. Trofymchuk O. M., Bidiuk P. I., Prosiankina-Zharova T. I., Terentiev O. M. Decision support systems for modelling, forecasting and risk estimation: monography. Riga: LAP LAMBERT Academic Publishing. 2019. 176 p. ISBN: 978-3-659-61214-5

Статті у періодичних виданнях, включених до категорії «А» Переліку наукових фахових видань України

2. Trofymchuk O.M., Bidiuk P.I., Prosiankina-Zharova T.I., Terentiev O.M. Bayesian data analysis in modeling and forecasting nonlinear nonstationary financial and economic processes. *Проблеми керування та інформатики*, 2023. № 4. С. 71-83. <https://doi.org/10.34229/1028-0979-2023-4-6>

Статті у наукових фахових виданнях України

3. Терентьев О. М., Кириченко В. Е., Связинська Н. О., Просьянкина-Жарова Т. І. Прогнозування фінансових ризиків з використанням наївного і доповненого деревом класифікаторів на основі байєсівських мереж. *Наукові вісті НТУУ КПІ*. 2016. №2. С. 60-68. <https://doi.org/10.20535/1810-0546.2016.2.63882>

4. Терентьев О. М., Просьянкина-Жарова Т. І., Савастьянов В. В. Використання засобів текстової аналітики як інструменту оптимізації підтримки прийняття рішень у задачах розробки планів соціально-економічного розвитку України. *Реєстрація, зберігання і обробка даних*. 2016. Т. 18. № 3. С. 75-86. URL: http://nbuv.gov.ua/UJRN/rzod_2016_18_3_10

5. Терентьев О. М., Просьянкина-Жарова Т. І., Савастьянов В. В. Застосування когнітивного та ймовірнісного моделювання у задачах формування сценаріїв розвитку соціально-економічних систем. *Наукові вісті НТУУ КПІ*. 2016. №5. С. 37-47. <https://doi.org/10.20535/1810-0546.2016.5.79876>

6. Бідюк П. І., Терентьев О. М., Просьянкина-Жарова Т. І., Ефендієв В. В. Прогнозне моделювання нелінійних нестационарних процесів у рослинництві з використанням інструментів SAS Enterprise Miner. *Наукові вісті НТУУ КПІ*. 2017. №1. С. 24-36. <https://doi.org/10.20535/1810-0546.2017.1.87423>

7. Bidyuk P., Overmyer S., Prosyankina-Zharova T., Terentiev O. Methodology of Modeling and Forecasting Nonlinear Processes in Finances. *Наукові вісми HTУУ КІІІ*. 2018. №1. С. 15 – 25. <https://doi.org/10.20535/1810-0546.2018.1.120361>

8. Bidyuk P., Prosyankina-Zharova T., Terentiev O., Medvedeva M. Adaptive modelling for forecasting economic and financial risks under uncertainty in terms of the economic crisis and social threats. *Технологічний аудит та резерви виробництва*. 2018. №4. С. 24-36. <https://doi.org/10.15587/2312-8372.2018.135483>

9. Bidyuk P. I., Korshevnyuk L. O., Gozhyi O. P., Kalinina I. O. , Prosyankina-Zharova T. I., Terentiev O. M. Modeling and forecasting financial and economic processes with decision support system. *Наукові вісми КІІІ*. 2019. № 5-6. С. 7-17. <https://doi.org/10.20535/kpi-sn.2019.5-6.176835>

10. Bidyuk P.I., Prosyankina-Zharova T.I., Diakon V., Diakon D. Modeling and forecasting nonlinear nonstationary processes in financial sphere. *Technology Audit and Production Reserves*. 2023. Vol. 4. № 2(72), P. 37 – 46. <https://doi.org/10.15587/2706-5448.2023.286516>

11. Terentiev, O., Prosyankina-Zharova, T., Diakon, V., Manuilenco, R. Development of mathematical models to support decision-making regarding the functioning of critical infrastructure in the industry of energy supply. *Technology Audit and Production Reserves*, 2023. 6(2(74)), 44–49. <https://doi.org/10.15587/2706-5448.2023.293205>

12. Лисов Б., Гуськова В., Просянкіна-Жарова Т., Халигов А. Інформаційна технологія класифікації даних моніторингу систем і мереж передачі даних для виявлення кіберзагроз. *Вісник Національного технічного університету" ХПІ". Серія: Інформатика і моделювання*. 2025. №1. Вип. 1(13). С. 128-147. <https://doi.org/10.20998/2411-0558.2025.01.09>

13. Просянкіна-Жарова Т. Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком. Екологічна безпека та природокористування, 2025. №54(2), С. 164–184. <https://doi.org/10.32347/2411-4049.2025.2.164-184>

Статті у наукових періодичних виданнях інших держав

14. Prosyankina-Zharova T., Terentiev O., Bidiuk P., Makukha M. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and forecasting. *Journal of Mathematics and System Science*. 2016. №6. P. 112-122. <https://doi.org/10.17265/2159-5291/2016.03.003> (CIIA, Index of Copurnicus)

Статті у закордонних виданнях, індексованих у базі даних Scopus

15. Bidiuk P. Modelling Nonlinear Nonstationary Processes in Macroeconomy and Finances/ Bidiuk P., Prosyankina-Zharova T., Terentiev O. / *Advances in Computer Science for Engineering and Education. ICCSEEA 2018. Advances in Intelligent Systems and Computing* / Ed. Hu Z., Petoukhov S., Dychka I., He M. (eds). Vol. 754. Cham: Springer, 2019. P.735-745. https://doi.org/10.1007/978-3-319-91008-6_72 (**Scopus Q3**). ISSN 2194-5357

16. Bidiuk P. I., Prosyankina-Zharova T. I., Terentiev O. M., Lakhno V. A., Zhmud O. V. Intellectual technologies and decision support systems for the control of the economic and financial processes. *Journal of Theoretical and Applied Information Technology*. 2019. Vol. 97. No. 1. P. 71-87. (**Scopus Q3**) URL: <https://www.jatit.org/volumes/Vol97No1/7Vol97No1.pdf>

17. Terentiev O. M., Prosyankina-Zharova T. I., Lakhno V. A., Usatiuk Y. V. The features of the predictive computing modeling power system load in terms of reforming energy market. *Journal of Theoretical and Applied Information Technology*. 2020. Vol. 98. No. 2. P. 163-182. (**Scopus Q4**). URL: <https://www.jatit.org/volumes/Vol98No2/1Vol98No2.pdf>

18. Terentiev O., Prosyankina-Zharova T., Savastyanov V., Lakhno V., Kolmakova V. The features of building a portfolio of trading strategies using the SAS OPTMODEL procedure. *Computation*. 2021. 9(7). 77. <https://doi.org/10.3390/computation9070077> (**Scopus Q2**)

19. Sharova, T., Zemlianska, A., Sharov, S., Prosyankina-Zharova, T., Kavun, L. Independent Evaluation of Learning Outcomes as a Tool of the Internal Quality Assurance System. *TEM Journal*. 2023, 12(2), P. 700–709. <https://doi.org/10.18421/TEM122-13> (**Scopus Q3**)

20. Sharov S., Chakraborty S., Zemlianskyi A., Kolmakova V., Prosiyankina-Zharova T. Quantitative Analysis of Online Courses for Learning the JavaScript Programming Language. *International Journal on Advanced Science, Engineering and Information Technology*. 2024. Vol. 14, No. 4, pp. 1216-1225. <https://doi.org/10.18517/ijaseit.14.4.19976> (Scopus Q4)

Наукові праці, які засвідчують апробацію матеріалів дисертації

21. Prosyankina-Zharova T. I., Terentiev O. M., Bidyuk P. I., Gasanov A. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and forecasting: *On control and optimization with industrial applications (COIA-2015)*, Proceedings of the fifth international conference (Baku, 27-29 August 2015), Baku: Institute of applied mathematics BSU, 2015. P. 365-368. URL: http://coia-conf.org/upload/editor/files/COIA_2015.pdf

22. Терент'єв О.М., Св'язінська Н.О., Прос'янкіна-Жарова Т.І. Візуалізація типології регіонів України із використанням система SAS Enterprise Guide 7.1. *Геоінформаційні системи і комп'ютерні технології еколого-економічного моніторингу*: матеріали міжнар. наук.-техн. конф. (Дніпропетровськ, 13–15 квітня 2016 р. Дніпропетровськ: ГВУЗ «НГУ» МОН України, 2016. С. 21-25. URL: <https://ela.kpi.ua/handle/123456789/19565>

23. Text mining analysis of agriculture internet sources using SAS software / Prosyankina-Zharova T. I., Terentiev O. M., Bidyuk P. I., Sviazinska N. O., Kyrychenko V. E. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2016)*: матеріали міжнар. конф. (м. Залізний порт, 24-28 трав. 2016 р.). Херсон: ХНТУ, 2016. С. 244-246. URL: <https://ela.kpi.ua/handle/123456789/19575>

24. Ivanova Y. V., Terentiev O. N., Korshevnyuk L. O., Prosyankina-Zharova T. I. Using modified logistic regression to increase customer response rate to marketing campaigns. *International conference on System Analysis and Information Technology (SAIT 2016)*. Proceedings of the 18-th conference (Kyiv, May 30 – June 2). Kyiv: Institute for Applied System Analysis, National Technical University of Ukraine

“KPI”, 2016. С. 304-305. URL: http://sait.kpi.ua/media/filer_public/73/32/7332a68e-e93b-4c57-a3c8-66f11ee074cd/sait2016ebook.pdf

25. Terentiev O.M., Prosyankina-Zharova T. I. Development of scoring systems for decision-making in crediting individuals and legal entities. *Розвиток підприємницької діяльності інтеграції України в світове господарство та інформатизації основних напрямків поступу*. Умань : Видавець “Сочінський М.М.”, 2017. С. 184-187. URL: <https://ela.kpi.ua/handle/123456789/19626>.

26. Терентьев О. М., Просьянкина-Жарова Т. И. Застосування прогностного моделювання для дослідження нелінійних нестационарних процесів у рослинництві. *Розвиток підприємницької діяльності інтеграції України в світове господарство та інформатизації основних напрямків поступу*. Умань : Видавець “Сочінський М.М.”, 2017. С. 113-116. URL: <https://ela.kpi.ua/handle/123456789/19619>.

27. Бідюк П. І., Терентьев О. М., Просьянкина-Жарова Т. И. Методи заповнення пропусків даних у задачах прогностного моделювання соціально-економічних процесів. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2017)* : матеріали міжнар. конф. (м. Залізний порт, 22-26 трав. 2017 р.). Херсон : ПП Вишемирський В. С., 2017. С. 185-187. URL: <https://ela.kpi.ua/handle/123456789/19679>.

28. Бідюк П. І. Терентьев О. М., Просьянкина-Жарова Т. И., Савастьянов В. В. Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни. *International conference on System Analysis and Information Technology (SAIT 2017)*. Proceedings of the 19-th. conference (Kyiv, May 22 – 25. 2017).. Kyiv: National Technical University of Ukraine “KPI”, 2017. С. 253-255. URL: http://sait.kpi.ua/media/filer_public/38/b9/38b90ff9-322e-4190-b445-e58ff05759fc/sait2017ebook.pdf

29. Bidyuk P. I., Prosyankina-Zharova T.I., Terentiev O.M. Multi model adaptive forecasting using decision support system *Моделювання та прогнозування*

економічних процесів. Матеріали XI Науково-практичної конференції.(м. Київ, грудень 6-8 грудня 2017 р.). К.: КПІ ім. Ігоря Сікорського, 2017, С. 14-17. URL <https://ecocyber.fmm.kpi.ua/wp-content/uploads/2022/06/mpep2017.pdf>

30. Трофимчук О.М., Бідюк П. І., Просянкіна-Жарова Т. І., Терентьєв О. М. Адаптивне моделювання і прогнозування у системах підтримки прийняття рішень сільськогосподарського призначення. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях*: колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2017. С. 21-28. URL: <https://itgip.org/wp-content/uploads/2018/08/zbirka-maket-1.pdf>

31. Bidyuk, P. Terentiev O., Prosyankina-Zharova T. Dynamic processes forecasting and risk estimation under uncertainty using decision support systems. IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Proceedings of the Proceedings of the conference (Kyiv, 29 May-2 June 2017), Kyiv: Igor Sikorsky Kyiv Polytechnic Institute. 2017. P. 795-800. (Scopus) <https://doi.10.1109/UKRCON.2017.8100355>. ISBN:978-1-5090-3006-4

32. Trofymchuk O. M. Forecasting nonstationary processes in demography, ecology, economy and finances/O. M. Trofymchuk, P. I. Bidyuk, T. I. Prosyankina-Zharova, O. M. Terentiev *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях*: колективна монографія/ ред. С.О. Довгий]. К.: ТОВ “Видавництво Юстон”, 2018. С. 113-123. URL: <https://itgip.org/wp-content/uploads/2018/10/zbirka-maket-2018-1.pdf>

33. Трофимчук О.М., Бідюк П.І., Терентьєв О.М., Просянкіна-Жарова Т.І. Особливості прогнозного моделювання з можливістю картографування та інтелектуального аналізу геопросторових даних у еколого-економічних дослідженнях; XVIII-а Міжнародна науково-практична конференція: «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: актуальні питання», колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво

Юстон”, 2019. С. 18-26. URL: https://itgip.org/wp-content/uploads/2019/10/1_knyga_sajt.pdf

34. Шолохов О. В., Терентьев О. М., Просьянкина-Жарова Т. І. Підвищення ефективності соціальних комунікацій на основі аналізу інтернет-джерел засобами text mining. *Прикладні системи та технології в інформаційному суспільстві: матеріали міжнар. наук.-практ. конф.* (Київ, 30 вер. 2020 р.). К.: Київський нац. ун-т імені Тараса Шевченка, 2020. С. 241-245. URL: https://aistis.knu.ua/wp-content/uploads/2020/11/tezi_2020_.pdf

35. Trofymchuk O.M., Bidyuk P.I., Prosyankina-Zharova T.I., Terentiev O.M. Decision support system for modeling and forecasting ecological processes. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: колективна монографія / ред. С.О. Довгий.* К.: ТОВ “Видавництво Юстон”, 2020. С. 23-44. URL: https://itgip.org/wp-content/uploads/2020/10/zbirka_2020_1.pdf

36. Korbicz J., Bidyuk P., Kuznietsova N., Terentiev O., Prosyankina-Zharova T. Multivariate distribution model for financial risks management. *The 9th International Conference "Information Control Systems & Technologies"(ICST2020)* (Odessa, Sept. 24–26, 2020). CEUR Workshop Proceedings, 2020. P. 416-429. URL: <https://ceur-ws.org/Vol-2711/paper32.pdf> (**Scopus**).

37. Trofimchuk O., Prosyankina-Zharova T.; Bidiuk P., Terentiev O. Operational risk estimation using probabilistic model. *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)* (Tashkent, Oct. 07-09, 2020), P. 1-5. <https://doi.10.1109/AICT50176.2020.9368630> (**Scopus**) ISBN:978-1-7281-7386-3, ISSN:2472-8586. URL: <https://ieeexplore.ieee.org/document/9368630>

38. Korbicza J., Trofymchuk O., Bidiuk P., Kuznietsova N., Terentiev O., Prosyankina-Zharova T. Probabilistic Modeling of Risks of Different Origin. *Modern Information Technologies for Environmental Safety Management, Nature Management, Emergency Measures*, October 7, 2021, Kyiv, Ukraine 2021. P. 29-42. URL: <http://ceur-ws.org/Vol-3021> (**Scopus**)

39. Трофимчук О. М., Бідюк П. І., Терентьєв О. М., Просянкін-Жарова Т.І. Застосування методики аналізу подібності часових рядів для прогнозування процесів різних типів в умовах інформаційної невизначеності; XX-а Міжнародна науково-практична конференція: *«Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: виклики 2021 року»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2021. С. 19-25. URL: https://itgip.org/wp-content/uploads/2021/10/1_zbirka_2021.pdf

40. Терентьєв О. М., Просянкін-Жарова Т.І., Дякон Д. В. Застосування засобів опрацювання неструктурованих даних у задачах прогнозного моделювання; XX-а Міжнародна науково-практична конференція: *«Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: виклики 2021 року»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2021. С. 161-168. URL: https://itgip.org/wp-content/uploads/2021/10/1_zbirka_2021.pdf

41. Терентьєв О.М. Просянкін-Жарова Т.І. Шолохов О.В. Особливості застосування інструментів інтелектуального аналізу даних компанії SAS у торгівлі криптовалютою *Прикладні системи та технології в інформаційному суспільстві*: матеріали міжнар. наук.-практ. конф. (Київ, 30 вер. 2022 р.). К.: Київський нац. ун-т імені Тараса Шевченка, 2022. С. 200-210. URL: <https://aistis.knu.ua/wp-content/uploads/2022/10/AISTIS-2022.pdf>

42. Trofymchuk O. M., Bidyuk P.I., Prosiankina-Zharova T. I., Terentiev O. M. Bayesian data analysis in modeling and forecasting nonlinear nonstationary financial and economic processes; XXI-а Міжнародна науково-практична конференція: *«Інформаційно-комунікаційні технології та сталий розвиток»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2022. URL: https://itgip.org/wp-content/uploads/2022/11/tezy_6.docx

43. Дякон Д.В., Просянкін-Жарова Т.І. Проблеми застосування інформаційних технологій у міжвідомчій співпраці в сфері реагування на різні загрози й надзвичайні ситуації. *Інформаційно-комунікаційні технології та*

сталий розвиток, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2022. URL: https://itgip.org/wp-content/uploads/2022/11/tezy_32a.pdf

44. Терентьев О. М., Просянкін-Жарова Т. І. Опрацювання інформаційної невизначеності із застосуванням методики виявлення подібності процесів. *Виклики і загрози для критичної інфраструктури*. матеріали міжнар. Наук.-практ. конф. (м. Київ, 29-30 червня 2023 р.) URL: <https://conference.cyberspace.org.ua>

45. Халігов А.А., Просянкін-Жарова Т.І., Гуськова В. Г. Удосконалення інформаційної технології аналізу даних на порталі open budget *Інформаційно-комунікаційні технології для перемоги та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2023.. С. 73–76. URL: https://itgip.org/wp-content/uploads/2023/11/1_zbirka_08_11_23-1-1.pdf

46. Бідюк П. І., Терентьев О. М., Просянкін-Жарова Т.І. Інформаційна технологія побудови прогнозів функціонування об’єктів критичної інфраструктури в умовах кризової ситуації. *Інформаційно-комунікаційні технології для перемоги та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2023. С. 114–117. URL: https://itgip.org/wp-content/uploads/2023/11/1_zbirka_08_11_23-1-1.pdf

47. Вишняков В.Ю., Глуган Ф.В., Просянкін-Жарова Т. І. Інформаційна технологія визначення тематичних точок на підстильній поверхні Землі із використанням даних дистанційного зондування Землі, отриманих з різних джерел. *Математичне моделювання та інформаційно-комунікаційні технології для зміцнення та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2024. С. 126-129. https://itgip.org/wp-content/uploads/2024/11/2024-11-24_zbirka_all_07_11_2024_148x210.pdf

48. Халігов А.А., Просянкін-Жарова Т. І., Гуськова В. Г. Використання рекомендаційної системи для підтримки прийняття рішень у системі кібербезпеки об’єкта критичної інфраструктури *Математичне моделювання та інформаційно-комунікаційні технології для зміцнення та відновлення*.

колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2024. С. 149-154. https://itgip.org/wp-content/uploads/2024/11/2024-11-24_zbirka_all_07_11_2024_148x210.pdf

Навчальні посібники

49. Згуровський М. З., Бідюк П. І., Терентьєв О.М., Просянкіна-Жарова Т. І. Байєсівські мережі в системах підтримки прийняття рішень: навч. посіб. К: ТОВ «Видавниче підприємство «Едельвейс», 2015. 300 с. ISBN 978-966-2748-73-4. URL: <https://ela.kpi.ua/handle/123456789/19582>

ЗМІСТ

ВСТУП

РОЗДІЛ 1 СИСТЕМОЛОГІЧНИЙ АНАЛІЗ ПРОБЛЕМИ ДОСЛІДЖЕННЯ УПРАВЛІННЯ РЕГІОНАЛЬНИМ РОЗВИТКОМ

1.1	Регіон та громада як об'єкти регіонального розвитку	35
1.2	Методологічні проблеми дослідження регіонального розвитку	42
1.3	Невизначеності та ризики в управлінні регіональним розвитком	47
1.4	Проблеми інформатизації регіонального розвитку	58
1.5	Постановка задачі дослідження	64
	Висновки по розділу 1	66

РОЗДІЛ 2 ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ ЗАСАДИ ПОБУДОВИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ

2.1	Теоретичні засади побудови систем підтримки прийняття рішень	69
2.2	Особливості використання інформаційних технологій у системі підтримки прийняття рішень управління регіональним розвитком	74
2.3	Методологія застосування мережецентричного підходу для побудови інформаційних технологій управління регіональним розвитком	79
2.4	Моделі та методи виявлення загроз та реагування на ризики загрози в інформаційній технології прийняття рішень	83
	Висновки до розділу 2	100

РОЗДІЛ 3	ІНФОРМАЦІЙНО-АНАЛІТИЧНЕ ЗАБЕЗПЕЧЕННЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ	
3.1	Особливості аналітичного забезпечення регіонального рівня	102
3.2	Моделі та методи попередньої обробки даних	104
3.3	Методика застосування моделі різних типів для аналізу та моделювання процесів регіонального розвитку	117
3.4	Застосування багатомодельного підходу	126
3.5	Методика моніторингу результативності державної регіональної політики в умовах невизначеності	136
	Висновки до розділу 3	149
РОЗДІЛ 4	МОДЕЛІ ТА МЕТОДИ ПРОГНОЗУВАННЯ РЕГІОНАЛЬНОГО РОЗВИТКУ	
4.1	Методичні засади прогнозування соціально- економічних процесів	151
4.2	Застосування методів прогнозування на основі часових рядів	155
4.3	Методика оцінювання адекватності моделей та оцінок прогнозів	159
	Висновки до розділу 4	172
РОЗДІЛ 5	МЕТОДОЛОГІЯ ПОБУДОВИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ У СИСТЕМІ УПРАВЛІННЯ РЕГІОНАЛЬНИМ РОЗВИТКОМ	
5.1	Архітектурні рішення щодо створення інформаційної технології	174
5.2	Розроблення прототипу системи підтримки прийняття рішень, що використовує розроблені інформаційні технології	160

Висновки до розділу 5	210
РОЗДІЛ 6 ПРИКЛАДИ ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ	
6.1 Інформаційна технологія удосконалення формування місцевого бюджету	213
6.2 Застосування інформаційних технологій прогнозування нелінійних процесів в енергетиці	222
6.3 Прогнозування еколого-економічних складових регіонального розвитку	225
6.4 Інформаційна технологія класифікації об'єктів для цілей регіонального розвитку	231
6.5 Застосування ансамблів моделей для оптимізації виробництва регіону	247
6.6 Моделювання еколого-економічних показників	247
6.7 Дослідження ефективності управління розвитком локальних соціально-економічних систем – об'єднаних громад	274
Висновки до розділу 6	281
ВИСНОВКИ	282
Список використаних джерел	284
ДОДАТКИ	
Додаток А	
Додаток Б	
Додаток В	
Додаток Д	

СПИСОК СКОРОЧЕНЬ

MRE	- mira Minimum Recognition Error
MSBs	- mean squared biases – середньоквадратичне зміщення
NDVI	- Normalized Difference Vegetation Index модуль для вирішення задач лінійного та нелінійного
OPTMODEL	- програмування
PCA	- Principal Component Analysis, метод головних компонент partial least squares regression model - регресія часткових
PLS	- найменших квадратів
R2	- коефіцієнт детермінації
RMSE	- Root mean square error - середньоквадратична похибка
SAS Base	- мова програмування для середовища SAS
SAS EM	- SAS Enterprise Miner SAS Interactive Matrix Language – інтерактивна матрична
SAS/IML	- мова програмування у середовищі SAS
SBC	- критерій Шварца-Байєса
SEMMA	- Sample, Explore, Modify, Model, Assessment
SMA	- simple moving average - просте ковзне середнє
SSE	- sum square error - сумарна квадратична похибка
STEPWISE	- методика поетапного включення регресорів smoothed weight of evidence – метод згладженого значення
SWOE	- зваженої сукупності методи об'єднання рівнів категоріальних змінних на основі
Threshold	- значення порогу
USD	- United States Dollar - долар США
VIF	- Variance Inflation Factor Weight of Evidence – вимірює статистичну значущість
WOE	- кожного класу змінної
АКФ	- автокореляційна функція
АРКС	- авторегресія із ковзним середнім
БД	- база даних
ІТ	- інформаційна технологія
КСС	- кількість ступенів свободи
ЛФК	- логарифмічна функція корисності
БМ	- Мережа Байєса
ВВП	- валовий внутрішній продукт
ВРП	- валовий регіональний продукт
МГК	- метод головних компонент

НКФ	- нелінійна кореляційна функція
НФК	- нелінійна кореляційна функція
ОМД	- опис мінімальною довжиною
ОПР	- особа що приймає рішення
ПК	- персональний комп'ютер
СППР	- система підтримки прийняття рішень
СУБД	- система управління базами даних
ЧАКФ	- часткова автокореляційна функція
ЩРЙ	- щільності розподілу ймовірностей

ВСТУП

Актуальність теми дисертаційного дослідження. Регіональний розвиток визначено у ст. 1 Закону України «Про засади державної регіональної політики» як процес соціальних, економічних, екологічних, гуманітарних та інших позитивних змін у регіонах, триває. Реформа місцевого самоврядування та територіальної організації влади, децентралізації частини владних повноважень та фінансів - складові Державної стратегії регіонального розвитку до 2020, посилили інституційну та фінансову спроможність місцевого самоврядування, що забезпечило стійкість громад та регіонів в умовах військової агресії. Наступним кроком є реалізація Державної стратегії регіонального розвитку на 2021-2027 роки, яка передбачає наближення державної регіональної політики України до політики згуртованості - ключової політики Європейського Союзу, що визначає такі пріоритети як відновлення, забезпечення якості надання державних послуг, підтримки економічного зростання, забезпечення добробуту населення. Реалізація вказаних пріоритетів потребує розроблення відповідних механізмів, методів та інформаційних технологій підтримки прийняття рішень, які дозволяють на основі всебічного та детального аналізу наявних даних, розробити декілька варіантів рішень, з яких особа, що приймає рішення може вибрати те, що найкраще відповідає поставленій задачі.

Розвинуті у роботі теоретичні засади вирішення проблеми інформаційно-аналітичного забезпечення підтримки прийняття рішень в управлінні регіональним розвитком ґрунтується на результатах досліджень вітчизняних та закордонних вчених: П. І. Бідюка, Х. Вотсон, С. О. Довгого, О. Г. Наконечного, О. В. Нестеренка, С. К. Полумієнка, В. Є. Снитюка, Р. Спрага, О. М. Трофимчука (*технології та системи підтримки прийняття рішень в умовах невизначеності*); В. Я. Данилова, М. З. Згуровського, В. С. Михалевича, Н. Д. Панкратової (*теорія та прикладні методи системного аналізу*), Г. Бокса, А. Дженкінса, В. І. Іваненка, О.Г. Івахненка, М. Кендалла, О. М. Новікова, О. А. Павлова, О. М. Хіміча (*математичне моделювання і прогнозування складних процесів і систем*);

П. І. Андона (*методи інтелектуального аналізу даних*); Р. Пірсона, О. І. Міхальова, Ю. П. Зайченка (*методи нечіткого аналізу і моделювання*); Т. Байєса, В. Гілкса, С. Лауритсена, Дж. Перла, Д. В. Стефанишина (основи ймовірнісного підходу); В. О. Васяніна, О. П. Гожого, В. А. Лахна, Т. В. Нескородевої, С. Ф. Теленика (*інформаційні технології*), В. М. Глушкова, В. І. Литвиненка (*системи та методи штучного інтелекту*), М. Бері, С. Аггарвала, Х. До Прадо, О. В. Копійки, Д. В. Ланде (*інтелектуальний аналіз тексту*) та інших.

Незважаючи на наявність значної кількості досліджень щодо розвитку соціально-економічних систем, проблема розробки моделей, методів та інформаційних технологій підтримки прийняття рішень в управлінні регіональним розвитком, залишається невирішеною повністю. Зокрема, слід відзначити відсутність загальних підходів до аналізу та попередньої обробки структурованих і неструктурованих даних, що описують предметну область в умовах невизначеності. Також нерозв'язаною залишається й низка задач, пов'язаних із виявленням характеру досліджуваних процесів, розробкою універсальних методик побудови моделей та їх ансамблів для отримання прогнозів розвитку процесів, характерних для соціально-економічних систем регіонального рівня за різних сценаріїв.

Розв'язання вказаних задач неможливе без створення єдиної науково обґрунтованої інформаційної технології, яка б охоплювала весь комплекс задач побудови прогнозів високої якості, необхідних для прийняття ефективних управлінських рішень.

На підставі викладеного можна стверджувати, що проблема побудови інформаційних технологій підтримки прийняття рішень в управлінні регіональним розвитком, які забезпечують високу обчислювальну ефективність, є актуальною.

Важливою і актуальною науково-прикладною проблемою, що розв'язується в дисертаційній роботі є підвищення ефективності підтримки прийняття рішень управління регіональним розвитком в умовах

невизначеностей та ризиків різної природи, на основі застосування сучасних методів системного аналізу, моделювання, прогнозування та інформаційних технологій.

Зв'язок роботи з науковими програмами, планами й темами. Тема дисертаційної роботи повністю відповідає тематиці наукових досліджень Інституту телекомунікацій і глобального інформаційного простору НАН України. Результати, отримані в ході виконання дисертаційного дослідження знайшли відображення у науково-дослідних роботах: «Розробка та аналіз засобів теоретико-ігрового моделювання стратегій збалансованого технологічного розвитку територій» (номер державної реєстрації 0116U000796), «Розробка інформаційної технології моделювання і прогнозування розвитку соціально-еколого-економічних систем в умовах невизначеності, нестаціонарності та ризику» (номер державної реєстрації 0121U100132), «Розроблення інформаційної технології комп'ютерного моделювання розвитку соціально-економічних систем під впливом глобальних чинників» (номер державної реєстрації 0122U200504), «Математичне моделювання, методи та інформаційно-комунікаційні технології для забезпечення стійкості критичної інфраструктури. Розділ 1. Розроблення математичних моделей, методів та технологій підтримки прийняття рішень забезпечення кібербезпеки та логістики критичних інфраструктур» (номер державної реєстрації 0123U100856), «Розробка інформаційних технологій та інструментальних засобів моделювання і прогнозування розвитку територій в умовах децентралізації» (номер державної реєстрації 0121U109211), «Математичне моделювання і прогнозування функціонування та розвитку соціально-економічної і транспортної інфраструктури територій і громад» (0123U104136), «Математичне моделювання, інтелектуальний аналіз даних та штучний інтелект для забезпечення прийняття рішень у сфері національної безпеки і оборони» (номер державної реєстрації 0125U000258).

Мета і завдання дослідження. Метою дослідження є розроблення єдиної формалізованої методології використання моделей, методів і методик системних досліджень у галузі інформаційних технологій, призначених для застосування в

системах підтримки прийняття рішень в управлінні регіональним розвитком на основі мережецентричного підходу, багатомодельного і багатокритеріального аналізу та інтеграції різнотипних даних, що забезпечує підвищення ефективності державного управління та місцевого самоврядування в умовах невизначеності та ризику.

Для досягнення мети дослідження розв'язано такі завдання:

- виконати аналіз сучасних методів і підходів до математичного моделювання, інтелектуального аналізу даних та інформаційних технологій для підтримки прийняття рішень у системах управління регіональним розвитком, з урахуванням умов невизначеності, нестаціонарності та різної природи процесів;
- проаналізувати сучасний стан і методи управління розвитком громад та регіонів, оцінити їх ефективність в Україні та за кордоном, а також виявити ключові фактори, що визначають досягнення цілей розвитку;
- дослідити сучасні методи моделювання та прогнозування соціально-економічних і соціально-еколого-економічних процесів та визначити можливості їх застосування до умов розвитку регіонів України;
- побудувати математичні моделі та ансамблі моделей процесів, характерних для територіальних громад і регіонів, з метою прогнозування їх розвитку, оцінки сценаріїв розширеного відтворення та формування варіантів управлінських рішень;
- удосконалити метод прогнозування розвитку регіональних соціально-економічних систем шляхом урахування особливостей територіального розташування, функціонування об'єктів підприємницької діяльності та інфраструктури, а також структурно-параметричних невизначеностей моделей;
- розробити та апробувати моделі для оцінювання і прогнозування економічних та екологічних ризиків у ключових сферах діяльності громад, здійснити формально-математичний опис алгоритмів і методів генерації множини кандидатів-сценаріїв розвитку процесів та визначити заходи щодо зменшення ймовірності негативних наслідків;

- виконати оцінку адекватності побудованих моделей і на її основі сформулювати рекомендації щодо їх застосування для оцінювання стану та прогнозування розвитку громад і регіонів;
- розробити методику розв'язання багатокритеріальних задач підтримки прийняття рішень із використанням еволюційних процедур та генетичних алгоритмів, інтегрованих у загальну методологію;
- на основі системного підходу побудувати та реалізувати інформаційну технологію підтримки прийняття рішень в управлінні регіональним розвитком, придатну для впровадження у системи державного управління та місцевого самоврядування.

Об'єктом дослідження є процеси регіонального розвитку.

Предметом дослідження – моделі, методи, методики та інформаційні технології підтримки прийняття рішень щодо розвитку регіонів і громад в умовах невизначеності та ризику.

Методи дослідження. Теоретичну та методологічну основу дослідження складають *метод діалектичного пізнання* стану та особливостей перебігу процесів регіонального розвитку, *системний підхід* до визначення факторів і механізмів, що впливають на динаміку соціально-економічних систем регіонального та місцевого рівнів. *При узагальненні теоретичних та методологічних засад* математичного моделювання та прогнозування розвитку досліджуваних соціально-економічних – методи індукції та дедукції, аналізу і синтезу, аналогій і зіставлення, формалізації й моделювання, порівняння, дослідження часових рядів, когнітивного, кластерного та інтелектуального аналізу даних, сценарний підхід, елементи теорії прийняття рішень. Для аналізу і обробки інформації використані методи системного, інтелектуального та статистичного аналізу даних, багатокритеріального аналізу, експертного оцінювання. Для моделювання та прогнозування процесів регіонального розвитку – методи математичного, економетричного, ймовірнісно-статистичного моделювання, теорії подібності процесів, методи використання комбінованих прогнозів, дерев рішень та багатовимірних розподілів. Для

розроблення інформаційної технології використано методи синтезу, системного аналізу, підтримки прийняття рішень, багатокритеріального аналізу та байєсівського підходу; *для побудови практичних реалізацій* – методи та засоби проектування і реалізації систем підтримки прийняття рішень, прикладне програмування.

Впровадження комп'ютерних інформаційних систем, оснований на створеному інструментарії, у роботу органів державного управління, місцевого самоврядування, дорадчих структур стане в нагоді при ідентифікації об'єктів критичної інфраструктури, коротко- та середньостроковому прогнозуванні, формуванні перспективних планів, отриманні стратегічних екологічних оцінок щодо природокористування при реалізації складних інфраструктурних проєктів, транспортної інфраструктури, виробничої логістики, уточненні оцінок впливу на довкілля потенційно небезпечних для довкілля проєктів, а також можуть бути використані і органами місцевого самоврядування для оцінювання тенденцій розвитку певних громад та оцінюванні результативності урядування.

Наукова новизна дослідження визначається теоретичними і прикладними результатами, отриманими автором, а саме:

вперше розроблено :

- єдину формалізовану методологію підтримки прийняття рішень в управлінні регіональним розвитком, яка, на відміну від відомих охоплює повний цикл від збору та інтеграції різнорідних структурованих і неструктурованих даних до формування оптимального рішення, та базується на системному підході з використанням багатомодельного і багатокритеріального аналізу, інтероперабельної мережецентричної архітектури й інтелектуальних технологій опрацювання даних, що дозволяє забезпечити узгодженість стратегій розвитку та підвищити обґрунтованість управлінських рішень;
- метод підтримки прийняття рішень на основі мережецентричного підходу, який на відміну від чинних, забезпечує взаємодію органів державного управління та територіальних громад у розподіленому інформаційному середовищі з підтримкою автоматизованого обміну даними та адаптації

управлінських стратегій до зовнішніх і внутрішніх умов, що збільшить гнучкість і швидкість реагування на зміни ситуації;

- інформаційну технологію моніторингу державної регіональної політики, яка на відміну від наявних рішень, поєднала математичні моделі, методи інтелектуального аналізу даних і штучного інтелекту з рекомендаційною підсистемою для формування сценаріїв і варіантів управлінських рішень в умовах невизначеності, що дає змогу своєчасно виявляти відхилення та пропонувати оптимальні коригуючі заходи;
- методику комплексного оцінювання ризиків і невизначеностей для соціально-еколого-економічних систем, яка, на відміну від наявних, базується на багатокритеріальному виборі сценаріїв розвитку та забезпечує інтеграцію кількісних і якісних показників, включно з експертними оцінками, що забезпечує високоякісне прогнозування та зниження ймовірності негативних наслідків;
- методику автоматизованого реагування на загрози різної природи, яка на відміну від чинних, реалізує вибір оптимальної стратегії з множини можливих, з урахуванням прогнозної інформації та обмежень ресурсів, що дозволяє мінімізувати втрати та підвищити стійкість системи;
- підхід до побудови ансамблів математичних моделей регіонального розвитку, який на відміну від наявних має модифіковану комплексну структуру, збільшує адекватність та точність прогнозування сценаріїв зміни соціально-еколого-економічних систем, що дає змогу формувати надійні прогнози.

вдосконалено:

- метод синтезу інформаційних технологій на основі багатомодельного й багатокритеріального підходів, який, на відміну від відомих, передбачає системне використання методів збору, оброблення, інтегрування та аналізу різнотипної інформації, моделювання, прогнозування і багатокритеріального прийняття рішень, із можливістю автоматизованого інтегрування даних, адаптивним налаштуванням та урахуванням досліджуваних процесів, що збільшить ефективність комплексного аналізу;

- метод прогнозування на основі адаптивного підходу до моделювання та прогнозування з використанням регресійних, узагальнених лінійних і ймовірнісно-статистичних моделей у формі мереж Байєса, який, на відміну від наявних, враховує можливі структурно-параметричні невизначеності моделей і забезпечує коректний формальний опис причинно-наслідкових зв'язків, що дозволить підвищити точність прогнозів;
- інформаційну технологію, в основу якої покладено інтеграцію принципів системного аналізу, методів обробки та оцінювання якості даних і прогнозного моделювання з використанням нових математичних моделей та їх композицій, що збільшить достовірність результатів і адаптивність до різних предметних областей;
- інструментальні засоби для розв'язання прикладних завдань прогнозування та підтримки прийняття рішень у різних галузях, створені на основі запропонованих методів, моделей і технологій аналізу даних та експертних оцінок, що розширить можливості практичного використання результатів дослідження;
- критерії адекватності моделей та оцінки якості прогнозів, які, на відміну від відомих, забезпечать комплексну перевірку проміжних і кінцевих результатів моделювання, та підвищать достовірність прогнозів і якість підтримки прийняття рішень у процесах регіонального розвитку.

отримав подальший розвиток:

метод оцінювання ризиків і невизначеностей на основі процедури оцінювання варіантів розвитку ситуації, який, на відміну від існуючих, дозволяє враховувати невизначеності та ризики різних типів, що забезпечує підвищення точності оцінок і надійності остаточних управлінських рішень.

Практичне значення одержаних результатів полягає у створенні системної методології розроблення інформаційних технологій, призначених для підтримки прийняття рішень в управлінні регіональним розвитком в умовах невизначеності, зокрема, спричиненої військовим конфліктом. Результати теоретико-методологічного та емпіричного досліджень доведені до

використання в науково-практичній діяльності, а також апробовані та впроваджені за безпосередньою участю автора в практичну діяльність органів місцевого самоврядування, державного управління, підприємств та установ, що підтверджується актами впровадження Черкаської обласної державної адміністрації, Львівської обласної державної адміністрації, Уманської міської ради, Паланської територіальної громади, Васильковецької територіальної громади, ТОВ НВП «Агроресурссистеми», ТОВ «СП «Укрінтерм»», ГС «Асоціація інженерів сталих енергетичних технологій України», ТОВ «Алгоритм – Х», ТОВ «ГІС Солюшн». Результати використовуються й у навчальному процесі Інституту прикладного системного аналізу Національного технічного університету «Київський політехнічний інститут України імені Ігоря Сікорського».

Особистий внесок здобувача. Наукові положення, висновки, рекомендації, що представлені в дисертації – особистий доробок здобувача. Інформаційне дослідження за темою дисертації виконано автором особисто. У роботах, опублікованих у співавторстві, до особистого внеску здобувача належать: розроблення методології підтримки прийняття рішень в управлінні складовими регіонального розвитку [1, 4,5, 8, 13, 14, 21, 33, 34, 40, 45]; постановка задачі, адаптація методики моделювання до досліджуваних процесів [1, 3, 6, 7, 8, 10, 11, 15, 16, 18, 21, 23- 28, 32-35, 39, 41, 42, 46, 47, 49]; метод побудови ансамблів моделей регіонального розвитку [2, 4, 5, 6, 7 8, 18 26, 27, 29, 30, 32, 42]; методика комплексного оцінювання ризиків [1, 8, 12, 21, 31, 37, 38, 44, 46, 48]; критеріальна база для аналізу складових регіонального розвитку [2, 4, 5, 10, 11, 15, 16, 17, 19, 22, 23, 25, 26, 27, 32-36, 39, 44, 48]; методика автоматизованого реагування на загрози різних типів [12, 18, 21, 24, 25, 32, 37, 38, 39, 44, 46, 48]; інструментальні засоби для розв’язання прикладних задач прогнозування та підтримки прийняття рішень [2-11, 16, 18, 20, 23, 26-30, 33-35, 40, 43, 46-49]; архітектура системи підтримки прийняття рішень в управління регіональним розвитком [13, 14, 40, 43, 45]; інформаційна технологія підтримки прийняття рішень [2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 18, 23, 26, 30, 33, 34, 39, 40, 43, 45, 46, 48].

Апробація результатів дисертації. Результати дисертаційного дослідження були представлені більше, ніж на 22 міжнародних та всеукраїнських наукових та науково-практичних конференціях, а саме, 15-й міжнародній конференції «International conference on control and optimization with industrial applications» (COIA-2015) (м. Баку, 27-29 серпня 2015 р.); науково-технічній конференції «Геоінформаційні системи і комп'ютерні технології еколого-економічного моніторингу» (м. Дніпропетровськ, 13–15 квітня 2016 р.); міжнародній конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту» (ISDMCI 2016) (м. Залізний порт, 24-28 травня 2016 р.); 18-й міжнародній науково-технічній конференції «Системний аналіз та інформаційні технології» (SAIT 2016). (м. Київ, 30 травня – 2 червня 2016 р.); всеукраїнській науково-практичній конференції «Розвиток підприємницької діяльності інтеграції України в світове господарство та інформатизації основних напрямків поступу» (м. Умань, 26-28 квітня 2017); міжнародній конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту» (ISDMCI 2017): (м. Залізний порт, 22-26 травня 2017 р.); 18-й міжнародній науково-технічній конференції «Системний аналіз та інформаційні технології» (SAIT 2017), XI науково-практичній конференції «Моделювання та прогнозування економічних процесів» (м. Київ, 6-8 грудня 2017 р.); IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON) (м. Київ, 29 травня – 2 червня 2017 р.); 16-й міжнародній науково-практичній конференції «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях» (м. Київ, Пуща-Водиця, 3-4 жовтня 2017 р.); XVII Міжнародній науково-практичній конференції «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях» (м. Київ, 25-26 вересня 2018 р.); First International Conference on Computer Science, Engineering and Education Applications (ICCSEEA2018) (м. Київ, 29 травня – 2 червня 2017 р.); XVII Міжнародній науково-практичній конференції «Сучасні

інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: актуальні питання» (м. Київ, 01-02 жовтня 2019 р.); 14-й міжнародній науково-практичній конференції «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: актуальні питання» (м. Київ, 06-07 жовтня 2020 р.); 4-й міжнародній науково-практичній конференції «Прикладні системи та технології в інформаційному суспільстві» (AISTIS-2020) (м. Київ, 30 вересня 2020 р.); 9-й міжнародній конференції «Information Control Systems & Technologies» "(ICST2020) (м. Одеса, 24–26 вересня 2020 р.); 14th International Conference on Application of Information and Communication Technologies (AICT 2020) (м. Ташкент, 07-09 жовтня 2020 р.); XX міжнародній науково-практичній конференції «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: виклики 2021 року» (м. Київ, 04-08 жовтня 2021 р.); XXI Міжнародній науково-практичній конференції: «Інформаційно-комунікаційні технології та сталий розвиток» (м. Київ, 14-16 листопада 2022 р.); 5-й міжнародній науково-практичній конференції «Прикладні системи та технології в інформаційному суспільстві» (AISTIS-2022): (м. Київ, 30 вересня 2022 р.); міжнародній науково-практичній конференції «Виклики і загрози для критичної інфраструктури» (м. Київ, 29-30 червня 2023 р.); міжнародній науково-практичній конференції «Інформаційно-комунікаційні технології для перемоги та відновлення» (м. Київ, 14-15 листопада 2023 р.); міжнародній науково-практичній конференції «Математичне моделювання та інформаційно-комунікаційні технології для зміцнення та відновлення» (м. Київ, 12-13 листопада 2024 р.).

Публікації

За результатами досліджень опубліковано 48 наукових праць, у тому числі:
 - 1 монографія (розділ у колективній монографії);

- 12 статей у наукових періодичних виданнях, включених до Переліку наукових фахових видань України (в т.ч. 1 включених до категорії “А”);

- 6 статей у наукових періодичних виданнях інших держав з напрямку, з якого підготовлено дисертацію (в.т.ч. 6, що включені до міжнародних наукометричних баз SCOPUS та/або Web of Science Core Collection;

- 4 статті у виданнях, віднесених до першого — третього квартилів (Q1—Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports; з них, 1 стаття у виданні, віднесеному до другого (Q2), прирівнюється до трьох публікацій та 3 – у виданнях, віднесених до третього квартиля (Q3), у яких наукова публікація прирівнюється до двох наукових публікацій;

- 28 тез та доповідей на наукових конференціях.

Якість та кількість публікацій відповідають Порядку присудження та позбавлення наукового ступеня доктора наук.

Структура дисертаційної роботи. Дисертаційна робота складається зі вступу, реферату, переліку умовних позначень, шести основних розділів, висновків, списку використаних джерел і додатків. Робота викладена на 317 сторінках і містить 275 сторінок основної частини, 55 рисунків, 51 таблицю, і список використаних джерел із 320 найменувань. У додатках представлено, акти впроваджень результатів роботи, код додатку, результати розрахунків та рисунки, що не увійшли до основного тексту роботи.

РОЗДІЛ 1

СИСТЕМОЛОГІЧНИЙ АНАЛІЗ ПРОБЛЕМИ ІНФОРМАЦІЙНОГО ЗАБЕЗПЕЧЕННЯ УПРАВЛІННЯ РЕГІОНАЛЬНИМ РОЗВИТКОМ

1.1 Регіон та громада як об'єкти регіонального розвитку

Поняття «регіон» широко використовуване у різних сферах, таких, як економіка, географія, регіоналістика, державне управління, соціологія, природокористування, статистика, тощо. Взаємопроникнення сфер, призводить до неоднозначності тлумачення даного терміну. Особливо це стосується досліджень соціально-економічних систем, де поняття «регіон» є дискусійним.

Тому, для цілей даного дослідження важливим питанням є визначення змістовного наповнення даного поняття.

У роботах вітчизняних та закордонних науковців немає однозначного й чіткого підходу до визначення даного поняття. Одні зосереджую увагу на тому, що регіон є організатором виробничого процесу, складовою регіональної економіки – головної ланки системної організації суспільного відтворення [50, с. 79], інші, застосовуючи системний підхід, визначають регіон як складну п'ятикомпонентну систему, з її підсистемами, такими як «системоутворююча база, системообслуговуючий комплекс, екологія, населення, інфраструктура ринку» [51]. Регіон розглядається науковцями і як певна територія, і як сукупність природних ресурсів, населення і засобів виробництва, тощо [52-55]. У економічній науці регіон розглядається й з точки зору економічних відносин - визначають регіон-квазідержаву, регіон-квазікорпорація, регіон-інституційна одиниця, регіон-соціум [56]. З точки зору міжнародного поділу праці вирізняють наднаціональні та транснаціональні утворення [57].

Однак, більшість фахівців визначають «регіон» як територіальне утворення, що має власні риси, є цілісною соціальною, адміністративною одиницею – частиною держави (або декількох держав), виокремлену соціально-економічну систему, характеризується виробничо-господарською

спеціалізацією [58].

Використовуючи методологію системного аналізу, регіон розглядають як складну багатофункціональну систему, що має певний вектор розвитку [99]. Регіон як система характеризується адаптивністю, схильністю до самоорганізації, процеси, що відбуваються в них - складністю, значною кількістю горизонтальних та вертикальних зв'язків між рівнями та окремими елементами системи. Тобто, регіон можна визначити як складну соціально-еколого-економічну систему, що є складовою системи вищого рівня - держави. В свою чергу, регіон є сукупністю складних систем нижчого рівня – районів та територіальних громад, що входять до складу районів. Територіальні громади мають власні цілі розвитку, ресурсний потенціал, територію, визначену у певних межах, галузеву структуру економіки, екологію, соціальну сферу, тощо. Райони, об'єднуючи декілька громад, формують власні цілі розвитку, узгоджені з цілями громад, регіонів та держави. Регіон є відкритою системою, розвиток якої відбувається у взаємодії з іншими системами – іншими країнами, державою (системами вищого рівня), іншими регіонами та системами нижчого рівня – районами та територіальними громадами, які входять до складу районів у регіонах.

Європейська Хартія Регіоналізації [59] визначає «регіон» як «область, що з географічної точки зору утворює чітку єдність або однорідний комплекс областей, які являють собою закриту структуру. Їх населення характеризується через певні спільні елементи, що дозволяє зберігати і далі розвивати певну своєрідність, для того щоб прискорити культурний, соціальний та економічний прогрес» [59]. Асамблея європейських регіонів у Декларації щодо регіоналізму в Європі [60] визначає, що «Регіон – це територіальне утворення, яке сформоване в законодавчому порядку на рівні, що є безпосередньо нижчим після загальнодержавного, і яке наділене... самоврядуванням» [61].

Інтеграція України до світового співтовариства та асоціація до європейського економічного простору, збройна агресія проти України, створили передумови до переформатування інституціонального підґрунтя соціально-

економічної системи держави і, відповідно, внесли корективи й до основних засад державної регіональної політики. У вітчизняному законодавстві визначено, що «регіон - Автономна Республіка Крим, області, міста Київ та Севастополь» [62]. Також визначено й функціональний тип території - макрорегіон, мікрорегіон чи територіальна громада, «які характеризуються сукупністю спільних соціальних, просторових, екологічних, економічних, безпекових та інших особливостей» [62]. Макрорегіон - декілька регіонів чи їх частини, об'єднані за спільними ознаками, яким притаманні спільні проблеми розвитку, у межах яких реалізуються спеціальні для цієї території програми регіонального розвитку [62]. Мікрорегіон визначений як частина регіону, що характеризується територіальною цілісністю та особливостями розвитку, у межах якої реалізуються спеціальні для цієї території проекти регіонального розвитку. Слід відзначити, що рівню району у чинному варіанті типізації не приділено увагу. Саме регіон та територіальна громада визначені у Державній стратегії регіонального розвитку на 2021- 2027 роки об'єктами державної регіональної політики.

Тому, в роботі, для цілей дослідження регіонального розвитку розглядаються саме регіони (Автономна Республіка Крим, області, міста Київ та Севастополь») та громади – як складні соціально-еколого-економічні системи, що мають просторове розташування у межах, визначених чинним законодавством.

Чинним законодавством передбачено також визначення функціонального типу територій для планування відновлення та стимулювання розвитку, а саме території відновлення [63]; регіональні полюси зростання [63]; території з особливими умовами для розвитку [63]; території сталого розвитку [63]. Застосування даної типологізації регіонів і територій здійснюється в рамках територіально-орієнтованого та безпекового підходу до розроблення документів державної регіональної політики, що дозволить уніфікувати процеси підтримки прийняття рішень в управлінні регіональним розвитком, активніше застосовувати досвід регіонів (громад)-лідерів, виявляти проблеми регіонів

(громад)-аутсайдерів, типові ознаки сталого розвитку регіонів, тенденції, диспропорції у галузевій структурі, проблеми та переваги у інвестиційному забезпеченні, відтворенні регіонів та громад. Що також враховано у дослідженні - важливо дослідити динаміку регіонального розвитку однотипних за виробничою спеціалізацією, галузевою структурою, рівнем соціально-економічного розвитку, інвестиційним забезпеченням, розташованих територіально близько – таких, що можуть утворювати ефективні макрорегіони (мікрорегіони).

Розглядаючи регіон з точки зору системного аналізу - як соціально-еколого-економічну систему, важливо, як зазначають, фахівці [63-66] дослідити її ієрархію, розгалуженість зв'язків підпорядкування та зв'язків між об'єктами на кожному рівні, оскільки саме ці характеристики є важливими для проектування інформаційного середовища відповідної системи підтримки прийняття рішень.

З точки зору системного аналізу, регіони, райони, громади є складовими національної соціально-еколого-економічної системи - складної стратифікованої системи з багаторівневою ієрархічною організацією, рівні якої (макро- мезо- та мікрорівень) взаємодіють між собою, маючи окремі цілі та динаміку, але спрямовуються до загальної мети: забезпечення сталого функціонування та розвитку (рис. 1.1.).

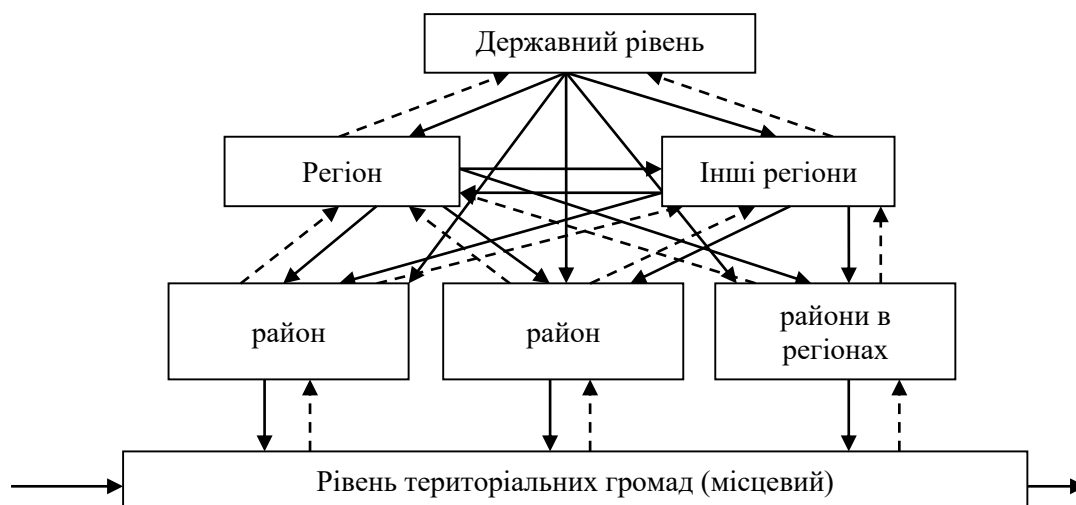


Рисунок 1.1 – Рівні системи управління соціально-еколого-економічної системи країни [63]

Розглядаючи соціально-еколого-економічну системи країни як правило виділяють державний, регіональний та локальний (або місцевий) рівень [644]. Кожен з рівнів характеризується повноваженнями, сукупністю ресурсів та економічних суб'єктів, взаємопов'язаних та взаємодіючих між собою у сфері виробництва, розподілу, обміну та споживання і утворюючих єдине ціле, а крім того, це специфічний соціум, що реалізує спільні інтереси населення країни, регіону, громади. Такий підхід відповідає представленому у чинному законодавстві визначенню «регіону». Тому у дослідженні використовуватиметься також і системний підхід, який базується на визначенні, регіону, представленому у чинному законодавстві.

Наближення України до європейського світовариста змінило й концептуальні засади регіональної політики, саму структуру регіональної системи управління. Чільне місце у регіональній системі посіли територіальні громади. Значною мірою це зумовлено тим, що в результаті децентралізації, держава делегувала більше повноважень територіальним громадам [65]. Зокрема, це питання утворення об'єднаних громад, формування та розподіл бюджетних коштів, визначення пріоритетів соціально-економічного розвитку, реалізації інвестиційних проектів, тощо. Маючи власні повноваження, громади, є об'єктами управління, які спільно вирішують і задачі регіонального розвитку, і власні, локальні задачі. Така модель управління відповідає мережецентричному підходу [66], який дозволяє забезпечувати гнучкість процесів управління складними системами, такими, як регіони. За такого підходу, кожна територіальна громада є окремим елементом регіональної системи, який може генерувати та отримувати інформацію, приймати рішення на основі отриманої інформації та ситуації, що складається, самоорганізовуватись, виконувати задачі регіонального розвитку в межах власних повноважень, відповідно до програм, планів, стратегій розвитку громад, районів та регіонів [67]. Тобто, фактично, регіональний розвиток відбуваються у розподіленому інформаційному середовищі, кожний вузол якого – громада, район має ситуаційну обізнаність, здатний до самоорганізації та прийняття рішень щодо використання наявних

ресурсів та забезпечення сталого функціонування.

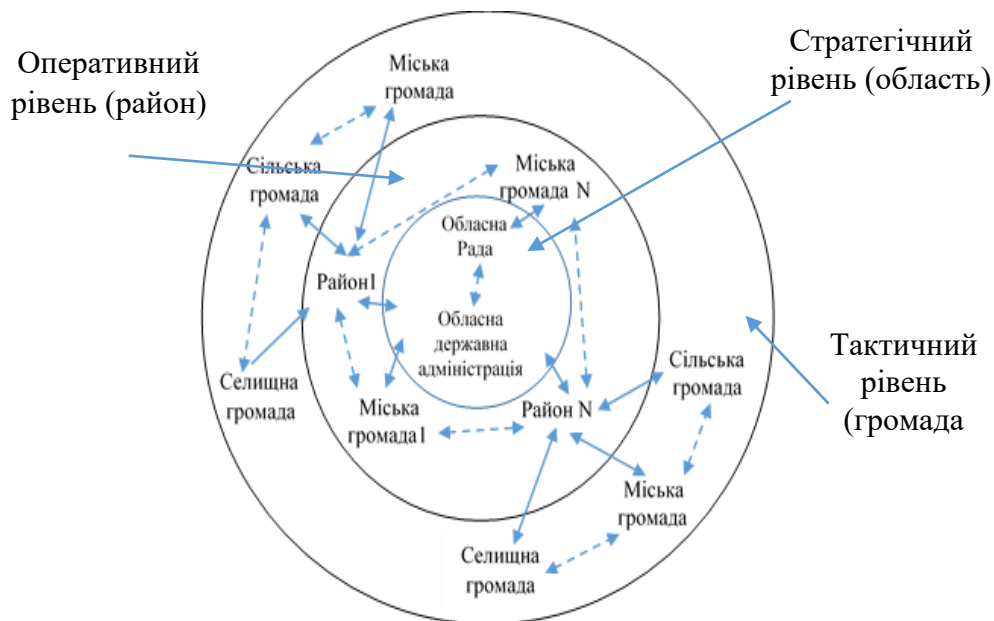


Рисунок 1.2 – Схема взаємодії об'єктів у системі управління регіональним розвитком [13]

Управління розвитком територіальних громад, в тому числі об'єднаних – це не лише питання вирішення нагальних питань, а й забезпечення збалансованого розвитку на перспективу. Їх розвиток відбувається не лише в умовах реформ, а й під впливом глобальних викликів, характеризується наявністю невизначеностей різного характеру. Питання забезпечення збалансованого розвитку, а можливо і розширеного відтворення в умовах докорінних змін державної регіональної політики є актуальним, але поки що не вирішеним.

Адже як у системі державного планування та прогнозування, так і під час оцінювання спроможності об'єднаних громад практично не застосовується сучасний інструментарій інтелектуальних технологій, прогнозного та сценарного моделювання, а процес прийняття рішень щодо перспектив розвитку громади очікуваних результатів реформ зосереджується на розрахунку

нормативних показників, фактично без аналізу потенційних ризиків та наслідків відповідних рішень.

Методика формування спроможних громад, затверджена постановою Кабінету Міністрів України від 24.01.2020 №34 [68] передбачає, що передумовами створення об'єднаних громад є оцінювання їх спроможності та розробка проекту перспективного плану формування територій громад області. Оцінювання спроможності громад здійснюється точково, із застосуванням показників на окрему дату, переважно застосовуючи методики якісного аналізу, такі як SWOT, PEST –аналізу, не враховуючи того, що процеси, які відбуваються у соціально-економічних системах, характеризуються нестаціонарністю є нелінійними, що значно ускладнює передбачення їх станів та забезпечення ефективного управління ними як на короткострокову перспективу, так і, особливо, на довгострокову. Відсутня також і чітка критична основа, що дозволяє виміряти результативність політики регіонального розвитку, зокрема, на рівні громад. Забезпечення ефективного використання ресурсів, формування інвестиційної політики адміністративно-територіальних одиниць, які входять (або планують входити) до спроможних територіальних громад, потребує ретельного дослідження та наукового обґрунтування. Зокрема, для реалізації інфраструктурних проектів, важливу роль відіграє і поточний стан системи, що аналізуються, її структура. А також, зв'язки в середині системи та динаміка процесів, що відбуваються. Важливим питання є й визначення взаємодії з іншими соціально-економічними системами, в тому числі і територіально віддаленими. Необхідно також враховувати макроекономічні наслідки децентралізації, стан навколишнього середовища, міжбюджетну взаємодію та суспільно-політичну ситуацію як окремих територій, так і в країні, територіальне розташування громади, а значить і потенційні можливості транскордонної взаємодії з сусідніми країнами, що існуюча методика не передбачає. Узгодження потребують як «горизонтальні», так і «вертикальні» взаємодії складових національної економіки та соціуму. Крім того, потребує удосконалення графічна частина перспективного плану, у якій, використовуючи геопросторові дані,

можливо унаочнити показники соціально-економічного розвитку, стану навколишнього середовища.

Отже, слід відзначити, що існуючий рівень впровадження інформаційних технологій в управлінні територіальними громадами як складними соціально-економічних системами не відповідає сучасним реаліям за якістю, надійністю, оперативністю, ступенем урахування різних типів невизначеностей і ризиків, обґрунтованістю стратегії розвитку територій на різні часові перерізи і за різних сценаріїв розвитку подій як у державі, так і в світі. Це пов'язано насамперед із недооціненістю переваг застосування інформаційних технологій для вирішення різних проблемно-орієнтованих завдань макроекономічного, соціально-економічного та соціально-еколого-економічного прогнозування, а з іншого боку - недостатньо розвиненим математичним забезпеченням, відсутністю адаптованих до проблемних ситуацій математичних моделей, методів та інформаційних технологій, що дозволяють вирішувати слабко структуровані та важко формалізовані задачі розвитку територіальних громад, розміщення виробництва, доступності об'єктів інфраструктури з урахуванням сучасних вимог, зокрема щодо збереження довкілля, раціонального природокористування, запобігання техногенних катастроф, поводження з об'єктами критичної інфраструктури.

1.2 Методологічні проблеми дослідження регіонального розвитку

Питанням регіонального розвитку, управління регіональним розвитком, прийняття рішень в управлінні регіональним розвитком, присвячено значну кількість робіт як вітчизняних, так і закордонних науковців та фахівців-практиків [69]. Але, як показало виконане інформаційне дослідження, незважаючи на те, що дане поняття відоме й часто використовуване, єдиного підходу до його визначення немає.

Поняття «розвиток» дуже часто вживають у таких сполученнях, як,

наприклад, «сталий розвиток», «економічний розвиток», «суспільний розвиток», «регіональний розвиток», «місцевий розвиток», тощо, які характеризують різні аспекти функціонування соціально-економічних систем та їх складових.

З філософської точки зору, поняття «розвиток», розглядається А. С. Капліною [70, с. 60] «як процес самопросування від низового рівня (простого) до вищого (складного), внаслідок якого відбувається розкриття і реалізація внутрішніх тенденцій та сутності явищ, що, у свою чергу, веде до виникнення нового і зумовлює будь-які зміни в різноманітних формах матерії». По-іншому визначення поняття «розвиток» наведено у «Філософському словнику» [71, с. 587]: «розвиток – специфіч. процес зміни, результатом якого є виникнення якісно нового..., поступальний процес сходження від нижчого до вищого, від простого до складного». У великому тлумачному словнику української мови [72, с. 1243] «розвиток» визначено як «процес, унаслідок якого відбувається зміна якості чого-небудь, перехід від одного якісного стану до іншого, вищого». Узагальнюючи різні погляди, слід відзначити, що багато вчених, [73, с. 59–61], визначають розвиток, перш за все як процес, що передбачає зміни у різних проявах.

Розвиток соціально-економічних систем характеризуються тим, що процеси розвитку вирізняються багатоваріантністю, мінливістю, переважно є нелінійними, їх перебіг відбувається в умовах невизначеності. Як зазначено у роботах [74-75], розвиток соціально-еколого-економічних систем є нелінійним динамічним процесом переходу системи з нестійкого у стійкий стан під впливом внутрішніх та зовнішніх факторів, які, зокрема, активізуючи механізми самоорганізації, спричиняють синергетичний ефект, що викликає позитивні зрушення та сприяє досягненню бажаного ефекту в економіці, соціальній сфері, екології (1.1) [4]:

$$\frac{dX}{dt} = F(S(t), R(t), P(t), C(t), D(t), L(t), I(t), E(t), Z(t)) , \quad (1.1)$$

де, $I(t)$ – інвестиції в економіку регіону; $R(t)$ – ресурсний потенціал економіки

регіону; $P(t)$ – природні ресурси, екологія, $D(t)$ – інформаційний потік; $C(t)$ – управляючий вплив; $L(t)$ – соціальна складова, $S(t)$ – стан економіки, $Z(t)$ – зовнішні фактори, $X(t)$ – очікуваний результат, $E(t)$ – ендogenous чинники, (t) – час [4].

Рушіями процесів регіонального розвитку є кібернетичні та синергетичні ефекти, як зазначають В. П. Мікловда та ін. [116, с. 7]. Наявність синергетичного ефекту, який, часто є суттєвим фактором економічного зростання – характерна особливість процесів розвитку регіонів та територіальних громад. Значна кількість дослідників відзначають, що регіони як системи, що самоорганізуються та саморозвиваються, а синергетичні ефекти є тими ендogenous чинниками, що створюють умови, які призводять до розвитку системи [117].

«Розвиток», як зазначають фахівці [118], «породжується протилежними процесами: дисипацією та оптимізацією – мінімізацією дисипації ресурсів системи», а «реалізовується через два типи його механізму – біфуркаційний та адаптаційний» [118]. Що також характерне й для соціально-еколого-економічної системи регіону.

«Регіональний розвиток» визначено в Законі України «Про засади державної регіональної політики» [103] як процес соціальних, економічних, екологічних, гуманітарних та інших позитивних змін у регіонах. Однак, цільові настанови, чітке означення того, що є критерієм позитивних змін в регіоні в даному визначенні не вказані. Тому, в рамках даного дослідження «регіональний розвиток» визначений як процес соціальних, економічних, екологічних, гуманітарних та інших позитивних змін у регіоні для забезпечення цілей сталого розвитку, розширеного відтворення та соціальних гарантій населенню.

«Державна стратегія регіонального розвитку на 2021-2027 рр» [103], розроблена відповідно до «Цілей сталого розвитку України до 2030 року» [119] в свою чергу, передбачає, що стратегічне планування має враховувати ключові виклики, які впливають на людину, інфраструктуру, безпекове середовище, економіку та навколишнє природне середовище, а також включати побудову культури партнерства та співробітництва, орієнтовану на взаємодію органів

державної влади, місцевого самоврядування та громадян.

Досягнення таких змін не можливе без дієвого та цілеспрямованого управління. Особливістю управління регіональним розвитком є те, що воно відбувається як з боку держави (вищий рівень), так і безпосередньо, на рівні регіону, і має бути скоординоване з пріоритетами територіальних громад. Тобто, мають бути поєднані централізоване управління, децентралізація на рівні територіальних громад та враховано специфіку регіонів. Що значною мірою відповідає мережецентричному підходу до управління [120].

Реалізація Державної стратегії регіонального розвитку на 2021-2027 рр. [3] передбачає розбудову багаторівневого урядування, адаптуючись до положень регіональної політики Європейського Союзу. Формування ефективного багаторівневого урядування потребує не лише організаційних заходів, а й відповідного інформаційного забезпечення, системного підходу до реалізації змін у існуючій структурі регіонального управління. Питанням багаторівневого управління присвячено багато робіт українських та закордонних науковців [121-125]. Впровадження багаторівневого урядування [126] передбачає перехід від суто ієрархічної структури, притаманної системі державного управління до системи урядування, яка передбачає налагодження системи партнерства між державою, бізнесом та суспільством. Таким чином змінюється система взаємодії – «управління» змінюється на «урядування» створюючи при цьому нові зв'язки та змінюючи характер взаємодії елементів регіональної системи. Ієрархічні зв'язки доповнюються горизонтальними – неієрархічними зв'язками, за якими відбувається взаємодія влади та суспільства. Відповідно, коригується існуюча модель організації управління регіональним розвитком. Зокрема, в результаті реформи децентралізації були закладені як фінансове підґрунтя, так і законодавча база місцевого самоврядування та державної регіональної політики, так і формування нового адміністративно-територіального устрою, залученню громадськості до публічного обговорення владних ініціатив, внесення пропозицій, оцінюванні влади. На даний час, реформування місцевого самоврядування та територіальної влади триває. Військовий стан із

запровадженням військових адміністрацій поки пригальмовує подальші зміни у системи регіонального управління, але у повоєнний час, очевидно, є перспективи до впровадження оновленої системи регіонального урядування.

Як свідчить вітчизняний досвід реформ та досвід країн Європейського співтовариства [127, с. 76], органи державного управління та місцеве самоврядування можуть активно співпрацювати, реалізовувати свої повноваження. Держава - як координатор регіонального розвитку, ефективно використовувати наявні аналітичні ресурси і управлінський потенціал для розроблення загальної концепції регіонального розвитку та заходів щодо її реалізації, регіони – та територіальні громади – а межах своїх повноважень (рис. 1.3).

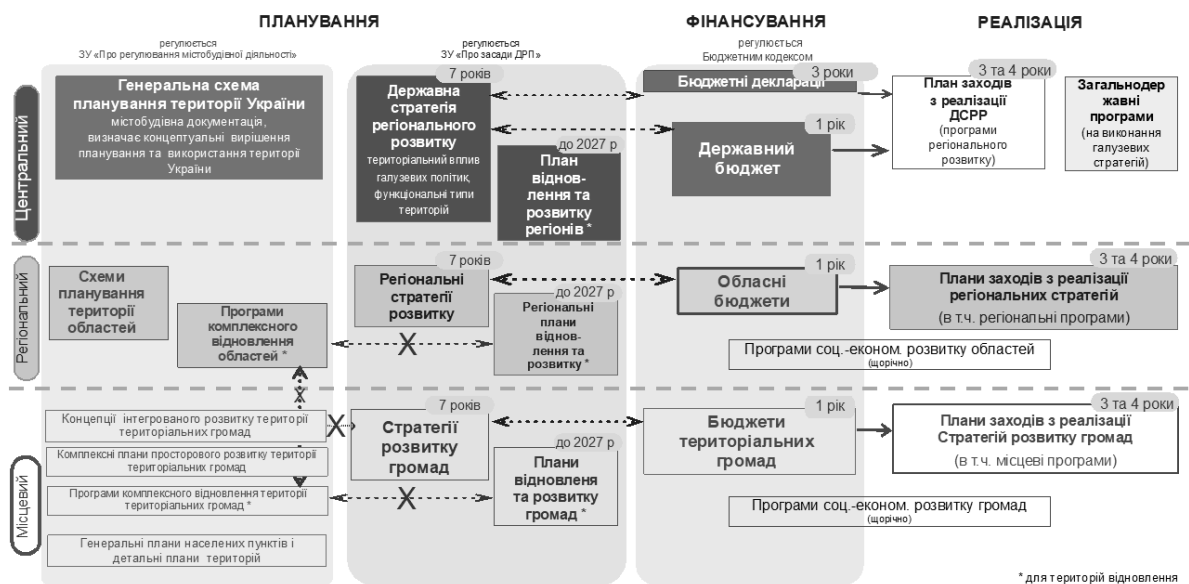


Рисунок 1.3 - Система планувальних документів державної регіональної політики <https://www.prostir.ua/?news=derzhavna-rehionalna-polityka-stan-spravna-traven-2024-roku> [128]

Отже, очевидно є потреба формування єдиного інформаційного забезпечення регіонального розвитку, яке відповідало б моделі європейського урядування, адаптуючи наявні моделі, методи, інформаційні технології до реалізації цілей регіонального розвитку.

1.3 Невизначеності та ризики в управлінні регіональним розвитком

Тривалі реформи, ситуація військового конфлікту створюють нові виклики, в тому числі й в управлінні регіональним розвитком. Зміни структури регіону, за рахунок укрупнення районів, реформування територіальних громад, ситуація військової агресії, кліматичні зміни, складна економічна ситуація, демографічні зміни спричинили ситуації, механізмів подолання яких на даний час ще не напрацьовано. Потрібен дієвий інструментарій для управлінців всіх рівнів, який би допомагав приймати «швидкі» рішення, аналізувати різні сценарії перебігу подій, прогнозувати майбутні зміни у регіоні та громаді, тощо. Ці проблеми є актуальними інтелектуальними завданнями, обтяженими як необхідністю залучення даних і знань з різних предметних областей, так і необхідністю подолання невизначеностей різних типів. А крім того, важливо, щоб розроблювані інформаційні технології стали звичайним засобом управлінської праці, їх застосування не потребувало б залучення фахівців-аналітиків.

Це ставить нові вимоги до розроблюваних інформаційних технологій, призначених для збору та аналізу інформації, генерування варіантів рішень, моделювання проблемних ситуацій та розроблення сценаріїв з оцінками ризиків і невизначеностей, вибором (побудовою) найкращого рішення.

Дослідження сучасних методів моделювання та прогнозування розвитку соціально-еколого-економічних систем виявило, що для відображення основних аспектів їх розвитку, проблема вибору ефективних рішень, може бути розв'язана за використання системного підходу, розглядаючи систему як складну ієрархічну структуру з множинними зв'язками (1.1) [моя стаття].

$$S_0 = S_1 \times S_2 \times \dots \times S_i \times \dots \times S_m \quad (1.1)$$

де S_i – i -й ієрархічний рівень, m – кількість ієрархічних рівнів [8].

$$S_i = \langle M_i, P_i, R_i, X_i, Y_i, f_i, \phi_i \rangle, \quad (1.2)$$

де, M_i, P_i, R_i – множина реальних об'єктів, суб'єктів та підсистем i -го ієрархічного рівня; X_i, Y_i – множина внутрішніх та зовнішніх параметрів системи i -го ієрархічного рівня та зовнішнього середовища ϕ_i, f_i – функціонали, що визначають взаємозв'язок відповідних параметрів на m рівнях у вигляді [8]:

$$\phi_i : X_i \rightarrow Y_i; \quad f_i : Y_i \rightarrow Y_{i-1};$$

Враховуючи, що соціально-економічна система розвивається в умовах невизначеностей, модель [1.2] доповнено урахування ризиків:

$$M = \{M_0(Y, I, P), M_E(X), M_{OE}, M_D(Q), M_{MO}, M_{ME}, M_U, A, M_H, M_{RS}\}, \quad (1.3)$$

де $M_0(Y, I, P)$ – ідентифікуюча модель системи, Y – ендогенні змінні, I – вектор керованих змінних, P – вектор ресурсів, $M_E(X)$ – модель навколишнього середовища, X – екзогенні змінні, M_{OE} – модель взаємодії об'єкта та навколишнього середовища, $M_D(Q)$ – модель поведінки системи, Q – збурюючі впливи, M_{MO} – модель зміни стану системи, M_{ME} – модель зміни стану навколишнього середовища, M_U – модель керуючої системи, A – правило вибору процесів зміни об'єкта, M_H – модель впливу дослідника на систему та результати дослідження, M_{RS} – модель системних ризиків [9].

Тоді, модель багатокритеріальної задачі прийняття рішень в умовах невизначеностей може бути представлена у вигляді [141] (1.4):

$$(M, Z, A, C, Cs, S, R), \quad (1.4)$$

де M – постановка задачі; Z – неконтрольовані невизначені фактори; A – множина можливих рішень; C – векторний критерій оцінювання рішень; Cs –

множина шкал критеріїв; S – множина інтегральних оцінок рішень; R – правило розв’язання задачі (прийняття рішень) [9].

Постановка задачі M у моделі (1.4) характеризує цілі особи, що приймає рішення та умови їх досягнення. Фактори Z зумовлюють наявність умов невизначеностей в задачі прийняття рішень, вони можуть бути представлені випадковими величинами. Множина допустимих рішень A становить сукупність рішень, що задовольняють наявним обмеженням з M , які розглядаються як можливі способи досягнення визначених цілей. Кожний варіант рішення оцінюється за сукупністю критеріїв, що представлені векторним критерієм $C = \{C_l\}$, $l = \overline{1, h}$. Для кожного критерію C_l задана шкала CS_l , яка є множиною упорядкованих оцінок. Шкали CS_l , $l = \overline{1, h}$ утворюють множину шкал критеріїв $CS = \{CS_l\}$. Кожному рішенню A_i з множини $A = \{A_i\}$, $i = \overline{1, n}$ надаються оцінки S_{il} за критеріями C_l , $l = \overline{1, h}$. Всі оцінки S_{il} , $i = \overline{1, n}$, $l = \overline{1, h}$ утворюють множину інтегральних оцінок рішень $S \subset Y$, де $Y = CS_1 \times CS_2 \times \dots \times CS_h$. R – правило розв’язання чи метод прийняття рішення, що можуть бути задані у вигляді аналітичного виразу, алгоритму чи вербального формулювання [1].

Вибір методу рішення R має вагоме значення в задачі прийняття рішень та суттєво впливає на якість остаточного результату. Однак, єдиного методологічного підходу до вирішення задач прийняття рішень в умовах невизначеностей немає [129]. Кожна задача є особливою і потребує остаточного висновку особи, що приймає рішення. Хоча є багато методів, алгоритмів обґрунтування управлінських рішень, однак, всі вони мають рекомендаційний характер, а вибір остаточного рішення завжди залишається за особою, що приймає рішення.

Формальна модель багатокритеріальної задачі прийняття рішення є прийнятною для дослідження та прогнозування розвитку процесів, що характерні для соціально-еколого-економічних систем. Для постановки задачі управління регіональним розвитком відповідно розробленим вимогам у роботі [1]

пропонується розширити модель (1.4), і ввести до її складу особу, що приймає рішення: $D = \{D_t\}$, $t = \overline{1, k}$.

Як зазначають [132-135], розв'язання задач прийняття рішення в умовах невизначеностей реалізується за наступною схемою:

- визначення типів невизначеностей, характерних для задачі;
- обрати метод урахування невизначеностей;
- застосування обраного методу врахування невизначеностей;
- постановка детермінованої багатокритеріальної задачі прийняття рішення;
- розв'язання багатокритеріальної задачі прийняття рішень.

Для оцінювання варіантів рішень з метою вибору оптимального варіанту, кожне рішення оцінюється за показниками (одним або кількома) для однієї або кількох моделей. Для опрацювання невизначеностей різних типів у задачах управління розвитком складних систем використовується множина методів для їх подолання залежно від типу невизначеності [136].

Невизначеність припускає наявність факторів, за яких результати впливу недетерміновані, а ступінь можливого впливу цих факторів на результати невідома [137-148].

Аналіз умов, за яких є невизначеності, можна виконувати на абстрактному теоретичному рівні або в умовах конкретної ситуації. Наприклад, з точки зору виявлення можливості побудови математичної моделі або на основі теорії інформації, де невизначеність виступає як характеристика ситуації вибору. Категорія невизначеності описується набором змінних параметрів, які описують різні групи невизначеностей, зокрема,: системну невизначеність, ситуативну, політичну, тощо. При вирішенні задач прийняття рішень в умовах наявності невизначеностей необхідно окреслити рівень аналізу і типи невизначеностей, що розглядаються.

Невизначеність – характерна особливість процесу прийняття рішень в управлінні [11]. Особа, що приймає рішення, завжди стикається з проблемою відсутності абсолютно повної інформації про об'єкт управління, розвиток подій, що супроводжують процес реалізації прийнятого рішення, наслідки.

Невизначеність станів об'єкту не є єдиною невизначеністю. Є й невизначеності, які пов'язані з суб'єктивізмом особи, що приймає рішення, її компетентністю, тощо. Крім того, не зважаючи на те, що в управлінні регіональним розвитком в більшості випадків цільові настанови визначаються у державних програмах, стратегіях, планах, невизначеність цілей та невизначеність критеріїв вибору рішень все ж присутні, оскільки кожний регіон, район чи громада мають особливі, притаманні лише їм умови, зв'язки, потреби. У багатьох реальних задачах складність прийняття рішення визначається насамперед числом альтернативних варіантів та кількістю і різноманіттям критеріїв оцінювання цих варіантів.

Застосовуючи методологію системного аналізу, фахівці визначають такі основні (найбільш поширені) типи невизначеностей [11, 136-148]:

- невизначеність цілей – невизначеність вибору цілей в багатокритеріальних задачах прийняття рішень [136-148];
- ситуаційна невизначеність – невизначеність впливу неконтрольованих факторів, що позначаються на процесах практичної діяльності [136-148];
- невизначеність природи процесу – відсутність достатніх знань ендгенних чинників, оточення та зовнішніх факторів [136-148];
- ненадійність очікувань – невизначеність розвитку певних подій у майбутньому [136-138];
- стратегічна невизначеність – невизначеність цілей і дій активного або пасивного партнера чи противника, так звана невизначеність конфліктів [136-148];
- інформаційна невизначеність – нечіткість та розпливчастість процесів і явищ та інформації про досліджувану систему, відсутність відомостей про достовірність інформації [136-148].

З точки зору моделювання та прогнозування [11, 136] розглядають такі типи невизначеностей, як:

- структурна невизначеність – невизначеність структури моделі досліджуваної системи [136-148];

- параметрична невизначеність – апріорна невизначеність параметрів моделі системи, складність оцінювання і аналізу якості параметрів моделі [136-148];
- статистична невизначеність – невизначеність статистичних даних, що переважно впливає з об'єму даних, відсутності (зміни методики) статистичних спостережень, наявності пропусків, імпульсних викидів тощо;
- методична невизначеність – невизначеність методу обробки даних чи методу розв'язання задачі [136-148];
- комбінаторна невизначеність – неможливість визначити всі варіанти розвитку подій [136-148].

В таких задачах управління регіональним розвитком часто стикаються з наступними невизначеностями, які виникають внаслідок сукупного впливу зовнішніх та внутрішніх чинників [11, 136, 138] :

- інформаційна невизначеність наявна в разі відсутності, неповноті і неточності інформації про певні аспекти розвитку досліджуваної системи чи процесів, про спроможність, рівень соціально-економічного розвитку громади, району, регіону, про програми, проекти, інвестиційні пропозиції, тощо;
- невизначеність цілей, зумовлюється, наприклад тим, що під час об'єднання громад не зрозуміло, які з них доцільно об'єднати, яким чином зробити це буде найраціональніше, вибору цілей при розв'язанні задачі формування громадського бюджету, тощо;
- ситуаційна невизначеність - невизначеність, що виникає наприклад, для задачі розподілу коштів державного бюджету на соціальні програми у різних районах або громадах області, коли інформації про ситуацію в окремих регіонах та громадах, про соціальну інфраструктуру окремих громад, інфляцію, обсяги виробництва продовольства, умов кредитування бізнесу, тощо; зміни у чинному законодавстві, військовою агресією, забезпечення інвестиційної привабливості регіону, змінами у фінансуванні окремих статей витрат місцевих бюджетів за рахунок надання субвенцій, тощо;
- стратегічна невизначеність, яка виявляється у неповноті і неточності інформації про цілі і наміри органів місцевого самоврядування у окремих

громадах стосовно соціально-економічного розвитку громад, об'єднання громад в межах району або декількох районів для вирішення спільних, наприклад, інфраструктурних проєктів, партнерів та конкурентів у галузях економіки регіону, щодо їх спроможності, банкрутств, інвестиційних проєктів з закордонними партнерами, ефективності нових виробництв, розміщення об'єктів переробної промисловості, підприємств, що виготовляють чи ремонтують сільськогосподарські машини, добрива, стосовно конкурентоспроможності їх товарів чи послуг, тощо.

Необхідно зазначити, що в процесі прийняття рішень в управлінні соціально-економічними системами, частіше всього, особа, що приймає рішення стикається з декількома різними видами невизначеностей, які в сукупності формують системну невизначеність [136-148].

Отже, особа, що приймає рішення, зазвичай стикається з проблемою розкриття системної невизначеності, що спричиняє ризикові ситуації, які не завжди вирішуються позитивно, тому питання розпізнавання настання ризикової ситуації не менш актуально.

Для розкриття невизначеностей різних типів розроблено багато підходів, що передбачають використання інструментів системного аналізу, математичного моделювання, морфологічного аналізу, тощо. Однак, єдиного універсального методу розкриття невизначеностей певного типу у задачах прийняття рішень немає. Це перш за все стосується слабоструктурованих проблем [11, 149] та задач з нечіткою вхідною постановкою задачі [150, 151].

Розкриття невизначеностей у задачах підтримки прийняття рішень в управлінні соціально-економічними системами можна визначити як задачу системно погодженого розкриття різнорідних невизначеностей на основі єдиних принципів, прийомів та критеріїв [11, 141].

Таким чином, ефективність формування управлінського рішення в задачах регіонального розвитку від набору методів, доступних особі, що приймає рішення, обраних способів розкриття і врахування наявних в задачі невизначеностей, а також від того, наскільки адекватно і в повній мірі за

використання цих методів можна відобразити досліджувану ситуацію.

Основні типи невизначностей та методи, використовувані для їх подолання представлені в таблиці 1.1.

Таблиця 1.1 – Основні невизначеності та методи їх розкриття [136-148]

№	Тип	Причини виникнення	Методи, використовувані для розкриття
1	Структурна	неможливо встановити всіх причинно-наслідкові зв'язки між змінними; наближені значення елементів структури моделі;	експертні методи; опис виявлених нелінійностей процесів за альтернативними аналітичними формами з подальшою оцінкою адекватності моделі та якості прогнозів; застосування статистичних тестів; застосування теорії перевірки гіпотез;
2	Статистична	Є похибки вимірів; стохастичні зовнішні збурення; мультиколінеарність; екстремальні значення; пропуски вимірів.	цифрові та оптимальні фільтри (фільтр Калмана); уточнення типів розподілів; метод головний компонент; теорія екстремальних значень; методи заповнення пропусків; просте усереднення
3	Параметрична	Не коректний вимір методу оцінювання; короткі вибірки.	забезпечення вибору альтернативних методів оцінювання параметрів; розмноження вибірок; методи Монте-Карло, ланцюги Маркова
4	Ймовірнісна	складність механізмів виникнення причинно-наслідкових зв'язків; відсутність детермінованості;	статистичні і динамічні мережі Байеса; марковські моделі; ймовірнісні фільтри; умовні розподіли;
5	Амплітудного типу	наявні змінні, що не вимірюються;	методи обробки нечіткої даних.
6	Ситуаційна	неможливо визначити всі причинно-наслідкових зв'язків	експертні методи когнітивне моделювання SWOT, PEST, SPACE аналіз

Для урахування невизначеностей застосовують такі методи, як [132, 134, 135]:

- метод коректив;
- аналіз чутливості;
- імовірнісний аналіз;
- дерева рішень;
- методи опису невизначених і неточних понять.

Узагальнення даних методів урахування і розкриття невизначеностей та аналіз їх придатності до застосування у задачах управління розвитком соціально-економічних систем, є складною задачею, яка потребує застосування сучасних інформаційних технологій у поєднанні із методологією системного аналізу.

Для урахування інформаційної, ситуаційної невизначеності, ненадійності очікувань та стратегічної, комбінаторної невизначеності в процесах, що протікають у соціально-економічних системах, де вплив людського фактору досить суттєвий, потрібні сучасні інформаційні технології, які б автоматизували процес формування варіантів рішень з урахуванням припущень щодо очікуваних значень. Вибір на користь впровадження інтелектуального та ймовірнісно-статистичного аналізу даних у інформаційних системах підтримки прийняття рішень ґрунтується на тому, що такий підхід до урахування невизначеностей позбавлений недоліків властивих іншим методам та постає придатним для опису та урахування системної невизначеності.

Більшість з проаналізованих методів розкриття та урахування невизначеностей в задачах прийняття рішень управління регіональним розвитком мають як переваги, так і недоліки, тому вони не можуть бути універсальними. Фахівці рекомендують використовувати комбінування різних методів, що надає можливість отримувати додаткові висновки щодо допустимих рішень.

Слід зазначити, що крім представлених підходів до врахування і розкриття певних невизначеностей при розв'язанні задач прийняття рішень, доцільним є дослідження розроблених методів багатокритеріального прийняття рішень в умовах невизначеностей. Це надає можливість враховувати наявні невизначеності та зводити задачі до детермінованих постановок. Такі методи або

їх елементи можуть бути використані в методології для усунення проблем, пов'язаних з факторами впливу системної невизначеності.

Детальне дослідження підходів до розв'язання багатокритеріальних задач та методів прийняття рішень, аналіз їх ефективності і можливостей застосування при розробці системної методології управління розвитком соціально-економічних систем, охоплює основні сучасні напрямки з прийняття багатокритеріальних рішень в умовах невизначеностей: методи узагальненого критерію, методи послідовної оптимізації, методи теорії ігор, передові підходи і методи штучного інтелекту [СТАТТЯ]. Оскільки більшість процесів в економіці, фінансах, екології, демографії, які відбуваються на рівні регіонів, характеризуються наявністю невизначеностей різних типів. Тому, для вирішення задач підтримки прийняття рішень в управлінні регіональним розвитком пропонується застосовувати системний підхід. Соціально-еколого-економічні процеси переважно є нелінійними, характеризуються наявністю нестационарності, потрібне удосконалення методології моделювання і прогнозування таких процесів на основі нових структур моделей, у яких поєднане застосування різних методів - статистичного, ймовірнісного (байєсівського), когнітивного, тощо, підходів подолання невизначеностей. Як представлено в роботах [наші статті] Таке об'єднання методів математичного, економетричного, економічного моделювання та системного підходу дає можливість побудувати моделі високого ступеня адекватності для різних соціально-еколого-економічних процесів і обчислити на їх основі оцінки прогнозів високої якості [НАШІ СТАТТІ].

Отже, для розв'язання задач управління регіональним розвитком, враховуючи всі основні аспекти проблеми вибору ефективних рішень, пропонується використовувати багатомодельний підхід, коли вибір рішень проводиться із опрацюванням комплексу різних, в тому числі й математичних моделей, інтелектуального аналізу даних. Сукупність різних вимог, які висуваються до рішень, призводить до багатокритеріальної постановки задачі прийняття рішень в управлінні регіональним розвитком, оскільки для більшості

вимог, подання їх у вигляді системи обмежень часто не ефективно через неповноту вхідних даних для такого подання [НАШИ СТАТТІ]. Це дозволяє визначити й вирішити проблеми управління розвитком складних соціально-еколого-економічних систем як завдання багатомодельного та багатокритеріального вибору ефективних рішень на множині математичних моделей. У загальному випадку кожне рішення оцінюється за одним або кількома показниками в рамках однієї або кількох моделей.

Для оцінювання ризиків у роботах [НАШИ СТАТТІ] запропоновано методику аналізу й оцінювання ризиків та невизначеностей у процесі моделювання ситуацій та оцінювання, а також для вибору варіантів їх розвитку. Методика полягає в тому, що під час аналізу й оцінювання ситуації на кожному етапі прийняття рішень визначається тип і величина ризиків, а також оцінюються і враховуються ступінь і тип невизначеності для вибору варіанта розвитку ситуації. Підтримуючи думку фахівців [154], у задачі оцінювання величини ризиків у задач прийняття рішень щодо регіонального розвитку доцільно використовувати теорію ситуаційного ризику, а для оцінювання інших типів ризиків використовувати відповідні міри ризиків [154].

Процедура врахування ризиків та невизначеностей, запропонована П. І. Бідюком та Н. В. Кузнецовою [154], може бути використана для побудови прогнозу процесів регіонального розвитку. Ця процедура є ітераційною та багатокроковою й використовується залежно від кількості етапів прийняття рішень й кількості ситуацій, які потрібно оцінити [154].

Отже, в сучасних умовах реформування регіональної системи управління, військової агресії та зміни економічної ситуації, комплексний підхід до урахування ризиків під час прийняття рішень в умовах невизначеності є важливою складовою дослідження процесів регіонального розвитку, має бути використаний у системах підтримки прийняття рішень органів державного управління та місцевого самоврядування. Оскільки, на відміну від бізнесових інформаційних систем, у системах такого типу, ця складова опрацьована слабо. Невизначеність зовнішнього середовища, зміна внутрішньої системи посилюють

рівень ризиків, а множинність взаємопов'язаних факторів зумовлюють необхідність застосування комплексних підходів до аналізу й прийняття ефективних управлінських рішень.

Тому, враховуючи досвід [наши статті] та [наши статті], в роботі пропонується удосконалена інтелектуальна технологія, розроблена на базі процедур системного аналізу, яка дозволяє не лише об'єктивно оцінити поточний стан системи, а й запропонувати варіанти її розвитку, виявляти критичні елементи та забезпечувати гнучкість управління в умовах невизначеності.

1.4 Проблеми інформатизації регіонального розвитку

Комп'ютерні моделі широко застосовуються у сфері інформаційно-аналітичної діяльності, дослідженнях соціально-економічних систем різних типів та рівнів ієрархії. Розроблені та активно використовуються такі моделі як EURACE (агент-орієнтована модель європейської економіки)[155], FuturICT 2.0 (модель, що включає агент-орієнтоване моделювання, big data, блокчейн, тощо, призначена для моделювання цифрової економіки) [156], УКРМАКРО (система економіко-математичних моделей розвитку економіки України) [157]. Є також і модельні комплекси, такі як Vensim 5.0 [158](має засоби побудови системних поточкових діаграм, виконання сценарних розрахунків та аналізу результатів моделювання), AnyLogic та NetLogic [159] (підтримують агентні моделі), AutoStat [160] (призначена для побудови моделей, особливістю є наявність процедур пошуку на основі еволюційних стратегій), OptQuest [161] (програма для моделювання, особливістю є використання інтелектуального пошуку наборів альтернатив, рішення щодо вибору сценарію приймається автоматично), OPTIMIZ [162] (програми для моделювання, що застосовують еволюційні стратегії та нейронні мережі); SimRunner2 [163] (група програм, призначених для моделювання, реалізовані еволюційні стратегії, генетичні алгоритми, тощо), WITNESS Optimizer [164] (програми для моделювання, їх особливістю є

наявність процедур реалізації алгоритму Simulated annealing, «пошуку із заборonoю»), тощо. Крім того, розроблені і використовуються для планування та прогнозування макроекономічних показників такі програмні засоби як додатки, створені на платформі INFORUM [165], проект ЛИНК [166], тощо.

Оскільки вартість програмного забезпечення, призначеного для побудови моделей та прогнозування є досить високою, широко застосовуються і власні розробки, виконані із використанням засобів MS Excel [167], мовами програмування C++ [168], Python [169], тощо.

Однак, в цих програмних засобах не враховуються специфічні особливості аналітичних задач, які призводять до неможливості їх безпосереднього застосування для розв'язання складних задач підтримки прийняття рішень

Міжнародні організації, такі, як, ООН [170] застосовують власні розробки, призначені для вирішення проблеми прогнозування сталого розвитку з урахуванням регіональної специфіки. Але складність розроблення аналітичного інструментарію, залишила це питання відкритим. ООН [170], навіть, відійшла від загальної проблеми сталого розвитку до проблеми реалізації його 17 цілей, через, певним чином, й теоретичну неможливість її вирішення.

Близькими до тематики розробки дослідженнями займаються найбільші аналітичні компанії та організації світу. Розроблені потужні аналітичні платформи, широко застосовувані у бізнесових структурах. Одними з ключових напрямів розроблення інформаційних технологій для підтримки прийняття рішень стає забезпечення комплексного підходу, що охоплює різні сфери, включаючи управління, охорону навколишнього середовища та життєдіяльність суспільства, що має забезпечити стабільність та підвищення добробуту населення та країни.

Система публічного управління активно реформується, адаптуючись до змін, що відбулись внаслідок реформи децентралізації. Здобутки системи публічного управління в Україні відзначені у аналітичному звіті Sigma monitoring reports [6] Зокрема, це цифрова трансформація публічного управління, розбудова системи електронного урядування.. Станом на початок

2024 року, за даними Sigma monitoring reports, Україна впевнено рухається у напрямку цифровізації. Впроваджено ключові механізми електронного урядування, що сприяло покращенню позицій України в Індексі розвитку електронного урядування з 82-го місця у 2018 році до 46-го місця у 2022 році. Позитивною є динаміка використання національної системи електронної взаємодії між державними інформаційними ресурсами, електронними реєстрами, інформаційними системами «Трембіта» [6] - кількість цифрових транзакцій зростає з 62,7 мільйона у третьому кварталі 2021 року до 541,3 мільйона у першому кварталі 2023 року. Кількість користувачів порталу «Дія» досягла близько 1,5 мільйона на день. В Україні діє понад 1100 Центрів надання адміністративних послуг, які пропонують стандартизований набір послуг громадянам та бізнесу [7]. У Збірнику інноваційних практик у державному управлінні та адмініструванні для сталого розвитку, виданому ООН [8] також, відзначено досягнення електронного урядування в Україні.

Більшість областей вже презентували програми інформатизації. Вже розроблені й впроваджуються заходи, передбачені у програмі інформатизації, працює система е-урядування. Основна увага приділена здійсненню моніторингу та оцінюванню розвитку регіонів і територіальних громад, е-урядуванню, обслуговуванню населення. На початку 2024 року презентовано тестову версію Єдиної геоінформаційної системи здійснення моніторингу та оцінювання розвитку регіонів і територіальних громад [5],

Налагодження роботи Центрів надання адміністративних послуг в територіальних громадах, значно покращило обслуговування населення, а також сприяло оперативній актуалізації інформації, що зберігається у державних реєстрах. Впроваджується Національна електронна система управління відбудовою DREAM [9], представлено тестову версію Геоінформаційної системи регіонального розвитку [5]. На регіональному рівні триває розроблення та прийняття регіональних програм інформатизації, запроваджено Єдину інформаційну систему обліку Національної програми інформатизації [10], розроблено Типовий проект регіональної програми, проекту, робіт з

інформатизації [11] та Типовий проект програми, проекту, робіт з інформатизації органу місцевого самоврядування [12]. Це, звичайно, потужні кроки в напрямку удосконалення інформаційного забезпечення регіонального розвитку, однак, питання розроблення інформаційних технологій, призначених до використання у системах підтримки прийняття рішень в органах державного управління та місцевого самоврядування на рівні регіону та територіальних громад, поки залишається не вирішеним

Але, мало уваги приділено питанням автоматизації процесів підтримки прийняття рішень, організації інформаційної взаємодії місцевого, регіонального та державного рівнів, узгодженню інформаційно-аналітичної роботи, розробленню єдиної концепції збору та накопичення інформації щодо соціально-економічного розвитку громад для забезпечення комплексного, системного підходу до планування, прогнозування, прийняття рішень та аналізу їх ефективності для забезпечення сталого регіонального розвитку. Обласні програми інформатизації розробляються виходячи з першочергових проблем регіонів, з урахуванням наявних фінансових та технічних можливостей щодо їх реалізації. Однак, за такого підходу, регіони матимуть технологічно різні платформи, різний набір використовуваного програмного забезпечення, організацію зберігання даних та знань, тощо. Все це значно знижує інтегровуваність інформаційних технологій, що впроваджуються, знижує потенційні можливості масштабування таких систем..

Складність створення аналітичного інструментарію [171], систем підтримки прийняття рішень для регіонального рівня, полягає у необхідності урахування значної кількості специфічних факторів. Зокрема особливостей життєвого циклу ієрархії «регіон – район- громади» як конкретної соціально-еколого-економічної системи, її структури, моделі управління, територіального розміщення, внутрішніх зв'язків, зв'язків з іншими системами, вплив суспільно-політичних чинників, зовнішніх та внутрішніх впливів різного характеру. А також передбачити проблемні ситуації, сценарії перебігу подій, можливі їх наслідки та потенційні ризики настання кризових ситуацій. Крім того, побудова

інформаційно-аналітичних систем для органів державного управління та місцевого самоврядування чітко врегульована на законодавчому рівні. Встановлено набір показників, які підлягають моніторингу та аналізу. Кожний регіон, як зазначається у «Програмі інформатизації» [172], має представити власну програму інформатизації, яка б охоплювала напрямки, пріоритетні саме для нього. Але, за такого підходу, втрачається системність. Регіональний рівень управління розглядається відокремлено, з перенесенням акцентів на вирішення точкових, внутрішніх, проблем конкретних громад регіону. Розроблення системи підтримки прийняття рішень управління регіональним розвитком в цьому випадку, здійснюється орієнтуючись на поточну ситуацію, розв'язання нагальних задач, а не виходячи з стратегічних пріоритетів. орієнтуючись на перспективи розвитку, що призводить до «пригальмовування» розвитку. Мало уваги приділяється аналітичній складовій, прогнозуванню, обґрунтуванню управлінських рішень. Наприклад, потреба регіонів та громад в інвестиціях часто планується виходячи з необхідності виконання окремих програм чи проектів, навіть привабливих на поточний момент, але таких, що не забезпечуватимуть ні спроможність територіальних громад на перспективу, ні подальший регіональний розвиток. Значними є ризики того, що реалізація програм інформатизації для регіонів та громад буде різною за концепцією, якістю реалізації, набором реалізованих функцій, не масштабованою та не відповідатиме в повному обсязі вимогам інтероперабельності.

Потребує вирішення й проблема уніфікації інформаційних технологій збору, оброблення та зберігання значних обсягів різномірної та неструктурованої, інформації, забезпечення її захисту та надійного передавання між вузлами системи. відповідно чинних вимог.

Тому, у роботі пропонується системне використання уніфікованого набору інформаційних технологій, об'єднаних у єдиній інформаційній системі регіонального розвитку, побудованій за мережецентричним принципом на основі інтероперабельності, застосування комплексної системи захисту інформації з підтвердженою відповідністю. Використання пропонованих в

роботі інформаційних технологій забезпечить [173]:

- узгодженість стратегій та програм розвитку регіонів та громад;
- сумісність технологічного, інформаційного, програмного забезпечення;
- системність моніторингу соціально-економічного розвитку регіону;
- об'єктивність інформації, що використовується у системі;
- оперативність та обґрунтованість рішень органів місцевого самоврядування та державного управління;
- доступність інформації для внутрішніх користувачів системи та населення а також інших зацікавлених сторін із розмежуванням доступу до неї залежно від типу користувачів та їх повноважень у системі [173].

Отже, цифрова платформа має відтворювати мережецентричну модель керування складною системою – регіоном (рис. 1.4).

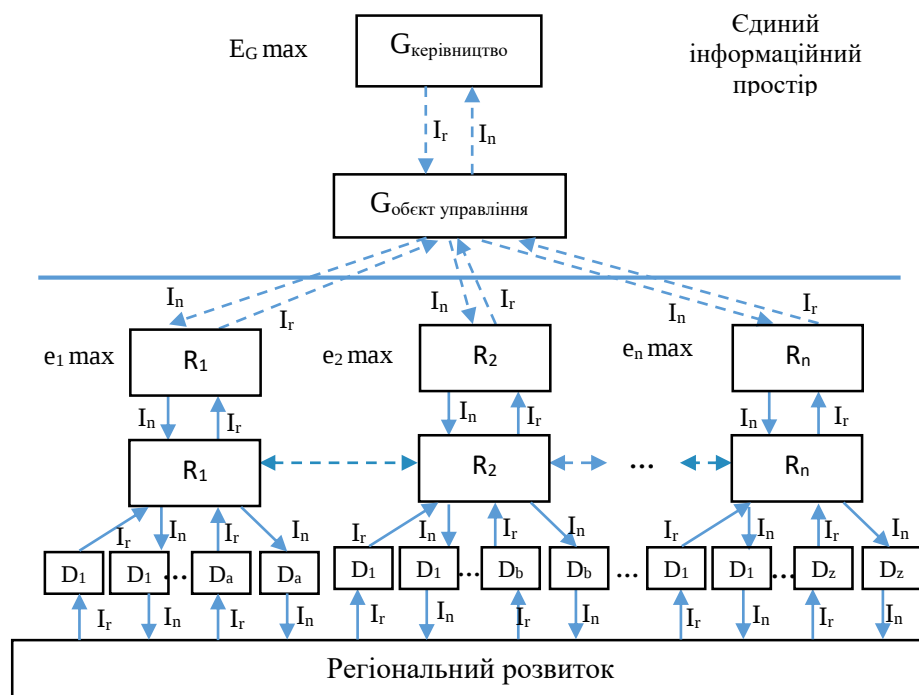


Рисунок 1.4 – Архітектура мережецентричної системи регіонального розвитку
[стаття одноосібна]

Найвищий рівень (G) – визначає цільову настанову системи – забезпечення визначених значень показників регіонального розвитку. На наступному – нижчому рівні ієрархії розглядається районний рівень управління (R , райони, кількість n) – підсистема, якою керують. Її цільова настанова (E) – забезпечення

сталого розвитку відповідно до найкращих показників. Підсистеми G та R обмінюються інформацією стосовно стратегії керування. Система вищого рівня (G) отримує інформацію про стан та функціонування системи нижчого рівня (R), коригує команди управління для об'єктів, що знаходяться на нижчому рівні (R), і, відповідно до зміни ситуації, коригує власні настанови. Найнижчий рівень – рівень територіальної громади (D, їх кількість різна для кожного району від 1 до z). Його ціль – забезпечення добробуту жителів громади. У процесі функціонування, системи, що знаходяться на одному рівні управління обмінюються інформацією, отримують настанови (In), повідомляють результати (Ir) коригують свою політику: система самоорганізується, адаптуючись до умов поточної ситуації, формуючи при цьому єдиний інформаційний простір [173].

Мережецентричне середовище утворюють фізичний, інформаційний та домен знань взаємодіють між собою створюючи синергетичний ефект.

Перевагою такої моделі є самостійність її ланок, оперативність прийняття рішень в умовах невизначеності, зокрема, системної, швидкої зміни ситуації, що особливо актуально в умовах війни. За такої побудови інформаційної системи, актуальною є проблема забезпечення інтероперабельності, оскільки мережецентрична архітектура передбачає, що всі елементи системи можуть обмінюватись інформацією «напрямую» один з одним. Однак, з іншого боку, така побудова системи дає можливість об'єднати різноманітні інформаційні системи територіальних громад і органів управління районного та обласного рівнів

1.5 Постановка задачі дослідження

Важливість чіткого та обґрунтованого формулювання цілей та завдань управління складними соціально-еколого-економічними системами, їхня складність і часові обмеження за яких приймаються рішення, потребують забезпечення дієвої комп'ютерної підтримки цього процесу. Створення інформаційно-аналітичних систем підтримки прийняття рішень, які забезпечили

б управлінців-практиків, експертів та дослідників сучасними засобами аналізу інформації, генерування варіантів рішень, моделювання проблемних ситуацій та сценаріїв, оцінками ризиків і невизначеностей, вибором і побудовою найкращого рішення є складними інтелектуальними завданнями, обтяженими як невизначеністю, необхідністю використання різномірної структурованої та неструктурованої інформації, необхідністю надійного зберігання даних і знань з різних предметних областей, в тому числі і за використання хмарних технологій.

Отже, виконаний аналіз інформатизації системи управління регіональним розвитком показав необхідність вирішення наступних проблем::

1. виконати аналіз сучасних методів і підходів до математичного моделювання, інтелектуального аналізу даних та інформаційних технологій для підтримки прийняття рішень у системах управління регіональним розвитком, з урахуванням умов невизначеності, нестаціонарності та різної природи процесів;
2. проаналізувати сучасний стан і методи управління розвитком громад та регіонів, оцінити їх ефективність в Україні та за кордоном, а також виявити ключові фактори, що визначають досягнення цілей розвитку;
3. дослідити сучасні методи моделювання та прогнозування соціально-економічних і соціально-еколого-економічних процесів та визначити можливості їх застосування до умов розвитку регіонів України;
4. побудувати математичні моделі та ансамблі моделей процесів, характерних для територіальних громад і регіонів, з метою прогнозування їх розвитку, оцінки сценаріїв розширеного відтворення та формування варіантів управлінських рішень;
5. удосконалити метод прогнозування розвитку регіональних соціально-економічних систем шляхом урахування особливостей територіального розташування, функціонування об'єктів підприємницької діяльності та інфраструктури, а також структурно-параметричних невизначеностей моделей;

6. розробити та апробувати моделі для оцінювання і прогнозування економічних та екологічних ризиків у ключових сферах діяльності громад, здійснити формально-математичний опис алгоритмів і методів генерації множини кандидатів-сценаріїв розвитку процесів та визначити заходи щодо зменшення ймовірності негативних наслідків;
7. виконати оцінку адекватності побудованих моделей і на її основі сформулювати рекомендації щодо їх застосування для оцінювання стану та прогнозування розвитку громад і регіонів;
8. розробити методику розв'язання багатокритеріальних задач підтримки прийняття рішень із використанням еволюційних процедур та генетичних алгоритмів, інтегрованих у загальну методологію;
9. на основі системного підходу побудувати та реалізувати інформаційну технологію підтримки прийняття рішень в управлінні регіональним розвитком, придатну для впровадження у системи державного управління та місцевого самоврядування.

Висновки до розділу 1

1. На основі аналізу літературних джерел визначено проблеми дослідження інформаційно-аналітичного забезпечення органів державного управління регіонального рівня та місцевого самоврядування.
2. Ввиконано аналіз сучасного стану та перспектив управління регіональним розвитком.
3. Проаналізовано підходи до визначення понять «регіон», «управління регіональним розвитком». Оскільки, визначення «регіону» в Україні та за кордоном суттєво відрізняються, то й «регіональний розвиток» у роботі досліджено з позицій вітчизняного законодавства, але враховуючи закордонний досвід розбудови економіки регіонів та місцевого самоврядування.
4. Зроблено висновок, що регіон є відкритою складною соціально-еколого-

економічною системою, розвиток якої відбувається у взаємодії з іншими системами – країнами, іншими регіонами та територіальними громадами, які входять до складу районів у регіонах. Фактично, регіон є сукупністю складних систем нижчого рівня – районів та територіальних громад, що входять до складу районів, мають власні цілі розвитку, ресурсний потенціал, територію у закріплених межах, галузеву структуру економіки, соціальну сферу, тощо, а з іншого боку разом, регіони є складовими системи вищого рівня – держави.

5. Встановлено, що особливістю управління регіональним розвитком є те, що воно відбувається як з боку держави (вищий рівень), так і безпосередньо, на рівні регіону, і має бути скоординована з пріоритетами територіальних громад, тобто мають бути поєднані централізоване управління, децентралізація на рівні територіальних громад та специфіка регіонів.
6. Встановлено, що єдиного підходу, адаптованого до потреб розв'язання задач управління регіональним розвитком немає.
7. Визначено, що опрацювання невизначеностей різних типів є необхідною складовою підтримки прийняття рішень у системі управління регіональним розвитком.
8. Встановлено, що в сучасних умовах системної невизначеності важливо не лише об'єктивно оцінити поточний стан системи, а й запропонувати варіанти її розвитку, виявляти критичні елементи та забезпечувати гнучкість управління в умовах невизначеності.
9. Встановлено, що потребує вирішення проблема уніфікації інформаційних технологій збору, оброблення та зберігання значних обсягів різномірної та неструктурованої, інформації, забезпечення її захисту та надійного передавання між вузлами системи. відповідно чинних вимог.
10. Визначено, що питання розроблення та системного використання сучасних аналітичних технологій у системі підтримки прийняття

рішеньв управлінні регіональним розвитком залишається не вирішеним.

11. Виконано постановку задач дослідження.

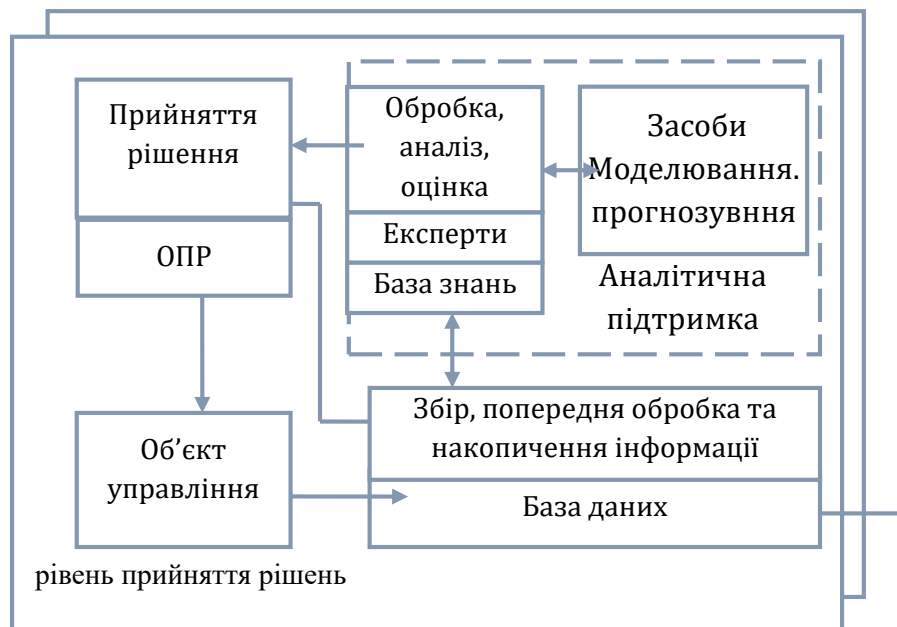
РОЗДІЛ 2

ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ ЗАСАДИ ПОБУДОВИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ

2.1 Теоретичні засади побудови системи підтримки прийняття рішень

Процес прийняття рішень -

Схематично, процес підтримки прийняття рішень [25] представлений на рис. 2.1.



Р

Розробленню систем підтримки прийняття рішень присвячено значну кількість робіт як вітчизняних так і закордонних фахівців []. існує багато визначень та класифікацій таких систем, підходів до їх проектування та реалізації. Найбільш поширеними є класифікації за призначенням, за концептуальною моделлю, за призначенням (завданнями, що вирішує), галузі застосування, тощо.

к

За класифікацією, наданою С. Альтером розрізняють такі типи систем, як

2

.

1

системи, орієнтовані на дані та системи, орієнтовані на моделі. За видом систем підтримки прийняття рішень, С. Альтмана за класифікує їх на 7 груп.

Таблиця 2.1 – Види систем підтримки прийняття рішень за класифікацією С. А.Альтмана

Вид	Призначення
Система накопичування файлів	Забезпечують особі, що приймає рішення оперативний доступ до певних елементів даних
Система аналізу даних	Створюють умов для аналізу даних
Система аналізу інформації	Надають користувачам інформацію використовуючи різні дані та моделі для їх оброблення
Розрахункові моделі або облікові та фінансові моделі	Покращення процесу планування через визначення наслідків дій на основі обчислювальних процедур
Репрезентативні та образні моделі	Генерація оцінки наслідків подій на основі імітаційних моделей
Оптимізаційні моделі	Вибір напрямів дій шляхом ідентифікації оптимальних рішень
Рекомендаційні моделі	Розроблення рекомендацій щодо напрямів дій на основі математичних перетворень та алгоритмів

Більш розширеною є класифікація Д. Павера, згідно якої, всі системи підтримки прийняття рішень поділяються на категорії за їх орієнтованістю на:

- дані (на основі сховищ даних, OLAP-системи, виконавчі інформаційні системи, територіальні),
- моделі,
- знання,
- документи,
- комунікаційні та групові (.

Та на три групи (за ознаками внутрішні-зовнішні користувачі, та спеціальні або загальні технології): втер- та інтра- організаційні; функціонально-специфічні та загального призначення, і побудовані на основі web-технологій. Серед СППР, орієнтованих на моделі

Крім вказаних класифікацій застосовують й інші, наприклад, за ознакою задежності собі, що приймає рішення – розрізняють незалежні (рішення приймається індивідуальні) та залежні, якщо рішення приймається колективом (або кількома особами)..

За типом архітектури, СППР,, як класифікують фахівці [27-30], розрізняють:

- текстово-орієнтовані;
- основані на використанні бази даних;
- на основі алгоритмічних процедур
- на основі правил
- гібридні [].

Не зважаючи на різноманіття класифікацій, спільною рисою СППР є те, що всі вони призначені для надання допомоги особі, що приймає рішення, а саме, вірно проаналізувати ситуацію, сформулювати варіанти рішень, визначити сценарії розвитку подій та оцінити можливі наслідки прийнятих рішень.

Основні підходи до проектування систем підтримки прийняття рішень, погоджуючись із класифікацією, наданою у [26], можна згрупувати наступним чином:

- інформаційний підхід;
- підхід, заснований на знаннях,
- інтелектуальний підхід

Прийняття рішень (рис. 2.2) в управлінні регіональним розвитком є складною задачею, яка відбувається в умовах складної суспільно-політичної та економічної ситуації, збройного конфлікту, активного реформування для адаптації до європейського економічного простору супроводжується наявністю невизначеностей та ризиків, вимагає застосування системного

міждисциплінарного підходу, розроблення комплексу інструментів аналізу та прогнозного моделювання складних систем, що ґрунтуються на використанні сучасних математичних методів, системному аналізі та прогнозуванні, комп'ютерних системах підтримки прийняття рішень.

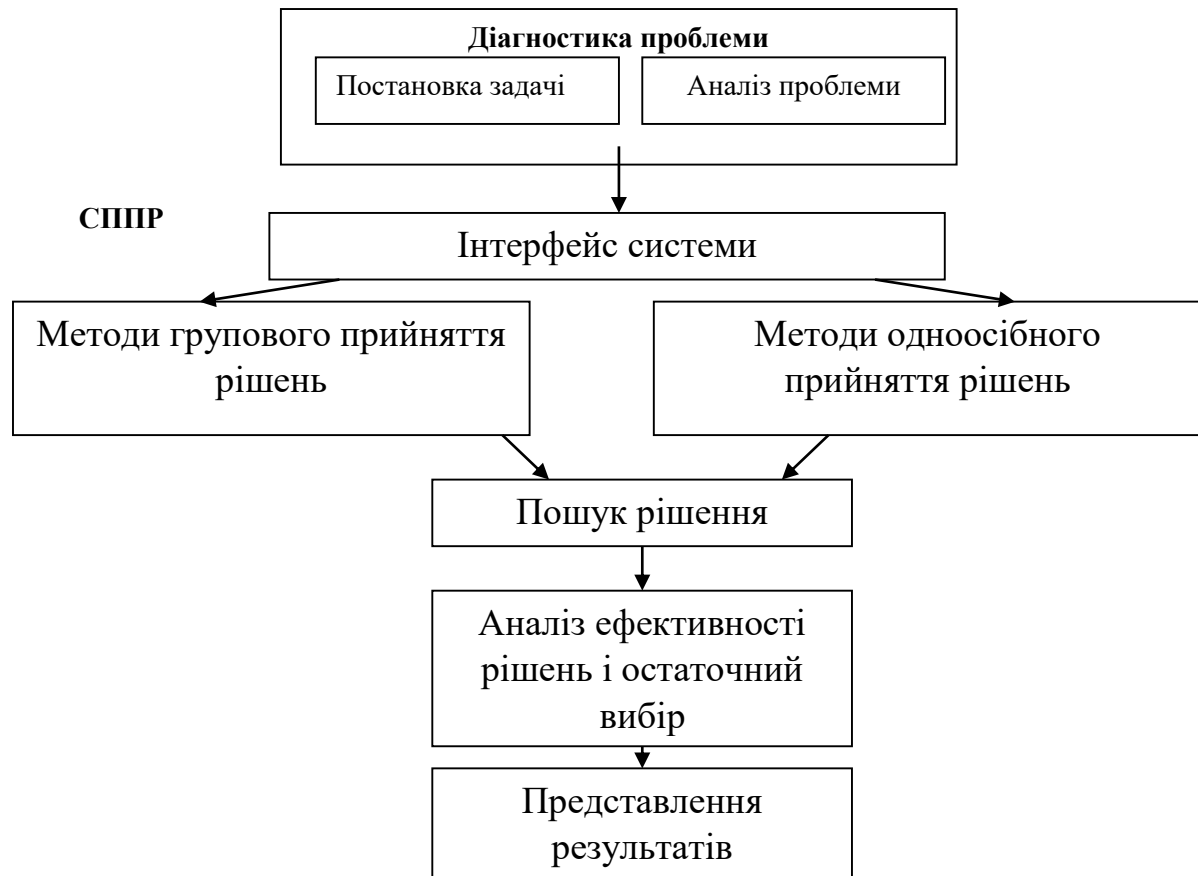


Рисунок 2.2 – Схема процесу прийняття рішень у СППР

Складність створення аналітичного інструментарію для підтримки прийняття рішень в управлінні регіональним розвитком пов'язана з необхідністю урахування значної кількості факторів, зокрема особливостей життєвого циклу соціально-еколого-економічної системи регіону, дослідження її структури, зв'язків, зовнішніх та внутрішніх впливів різного характеру. Потрібно також передбачити проблемні ситуації, сценарії розвитку подій, можливі наслідки, тощо. Потребує вирішення також проблема створення сучасних інформаційних технологій, які дозволяли б накопичувати та зберігати значні обсяги різномірних, структурованих та неструктурованих даних, що описують

різні аспекти регіонального розвитку, забезпечують швидку їх обробку та захист.

Для забезпечення сталого розвитку регіону, важливу роль відіграє не лише поточний стан його соціально-еколого-економічної системи, а й майбутня структура, динаміка процесів, зв'язки з іншими системами одного та різних рівнів. Важливо врахувати й вплив глобальних та макроекономічних чинників, міжбюджетну взаємодію, запровадження заходів державного регулювання економіки та суспільно-політичну ситуацію в регіонах, громадах, країні, екологічні фактори, загрози національній безпеці та ризики різних типів.

Тому, для цілей даного дослідження обрано методику підтримки прийняття рішень на основі використання набору моделей різних типів.

Крім того, при проектуванні системи підтримки прийняття рішень управління регіональним розвитком важливо врахувати, що для даного рівня важливими є не лише тактичні цілі - вирішення поточних проблем функціонування окремих закладів чи організацій. Рівень області має перш за все акцентувати увагу на розв'язанні стратегічних задач, визначати цільові пріоритети розвитку, координуючи їх із політикою держави та окреслюючи напрями для оперативного рівня прийняття рішень. Оперативний рівень – рівень районів (або в подальшому об'єднань районів або громад) має вирішувати задачі досягнення параметрів розвитку, передбачених на стратегічному рівні управління, скоординовуючи свої дії для досягнення стратегічних цілей та дії тактичного рівня регіонального управління. Не зважаючи на те, що система управління регіоном визначається чинним законодавством для всіх регіонів, регіони є складними слабкоструктурованими динамічними системами, тому кожний з них має свої особливості, які необхідно врахувати при побудові системи підтримки прийняття рішень. Крім того, всі суб'єкти регіону мають власні цілі, мають різне підпорядкування та зв'язки. Крім того, для регіонального рівня важливим є й територіального розташування, яке значною мірою визначає виробничу спеціалізацію регіону, ресурсний потенціал, тощо. Іншою проблемою, яка має бути вирішена у процесі розроблення системи - це проблема неоднорідності інформаційного середовища регіонів. Технічне,

технологічне різноманіття є суттєвою проблемою у створенні єдиного інформаційного простору. Крім того, є питання організації міжвідомчої взаємодії та співпраці держави з громадськістю, яка має активніше долучатись до вивчення проблем регіонального розвитку, особливо на рівні територіальних громад. Тому питання організаційного характеру в умовах децентралізованого прийняття рішень є чи не найскладнішими. Серед організаційних заходів необхідно передбачити засоби збору та оброблення даних щодо стану та динаміки системи, організаційні заходи щодо забезпечення моніторингу процесів регіонального розвитку виходячи з того, що дослідження регіонального розвитку розглядається з точки зору складових: соціальний розвиток, екологія, економічне зростання. Тому, у системі підтримки прийняття рішень управління регіональним розвитком необхідно приділити увагу інформаційному забезпеченню – створенню єдиного інформаційного середовища, орієнтованого на виконання аналітичних функцій, підтримки та координацію процесів прийняття рішень на всіх рівнях системи управління регіоном та забезпеченню зворотного зв'язку для аналізу та контролю за досягнутими результатами.

2.2 Особливості використання інформаційних технологій у системі підтримки прийняття рішень управління регіональним розвитком

Сукупність різних вимог, які висуваються до рішень, призводить до багатокритеріальності постановки задачі, оскільки для більшості вимог подання їх у вигляді системи обмежень не може бути визнано ефективним через неповноту вхідних даних для такого подання. Тобто, визначити й вирішити проблему ефективного управління регіональним розвитком можна як завдання багатомодельного та багатокритеріального вибору ефективних рішень на множині математичних моделей, що описують процеси регіонального розвитку. У загальному випадку кожне рішення оцінюється за одним або кількома показниками в рамках однієї або кількох моделей. Для опрацювання невизначеностей різних типів використовується множина методів для їх

подолання залежно від типу невизначеності. Соціально-еколого-економічна система регіону розглядається як складна ієрархічна структура з множинними зв'язками.

Окремою проблемою є узгодження взаємодії «горизонтальних» та «вертикальних» складових соціально-економічної системи країни та соціуму регіонів. Адже, сільський та міський соціуми регіонів мають суттєві відмінності, неврахування яких під час управління регіональним розвитком може призводити до суспільно-політичного напруження. Не зважаючи на значну кількість досліджень, відсутні адаптовані до реальних проблемних ситуацій математичні моделі, методи прогнозування та інші інструменти, що дозволяють розв'язувати слабко структуровані та важко формалізовані задачі управління розвитком територіальних громад у контексті реалізації задач регіонального розвитку.

Коло проблемних питань в сфері прийняття рішень в управлінні регіональним розвитком не обмежується наведеними. Їх вирішення напряму пов'язане з розробленням та впровадженням інформаційних технологій, спрямованих на розв'язання різних проблемно-орієнтованих завдань розвитку регіонів та територіальних громад на основі макроекономічного, соціально-економічного та соціально-еколого-економічного прогнозування. Існуючий рівень інформаційних технологій в управлінні регіоном і територіальними громадами, як складними ієрархічними просторово-розподіленими системами, не відповідає сучасним реаліям за якістю, надійністю, оперативністю, ступенем урахування різних типів невизначеностей і ризиків, обґрунтованістю стратегії розвитку територій на різні часові перерізи і за різних сценаріїв розвитку подій.

Тому, для вирішення задач розроблення інформаційних технологій прийняття рішень в управлінні регіональним розвитком запропоновано системну методологію, яка вирізається поєднанням рівних методів збору, оброблення, аналізу даних, когнітивного, математичного моделювання, сценарного аналізу, бізнес-аналізу, підтримки прийняття рішень. Загальна (спрощена) схема застосування інформаційних технологій для підтримки прийняття рішень в управлінні регіональним розвитком (рис.2.3).

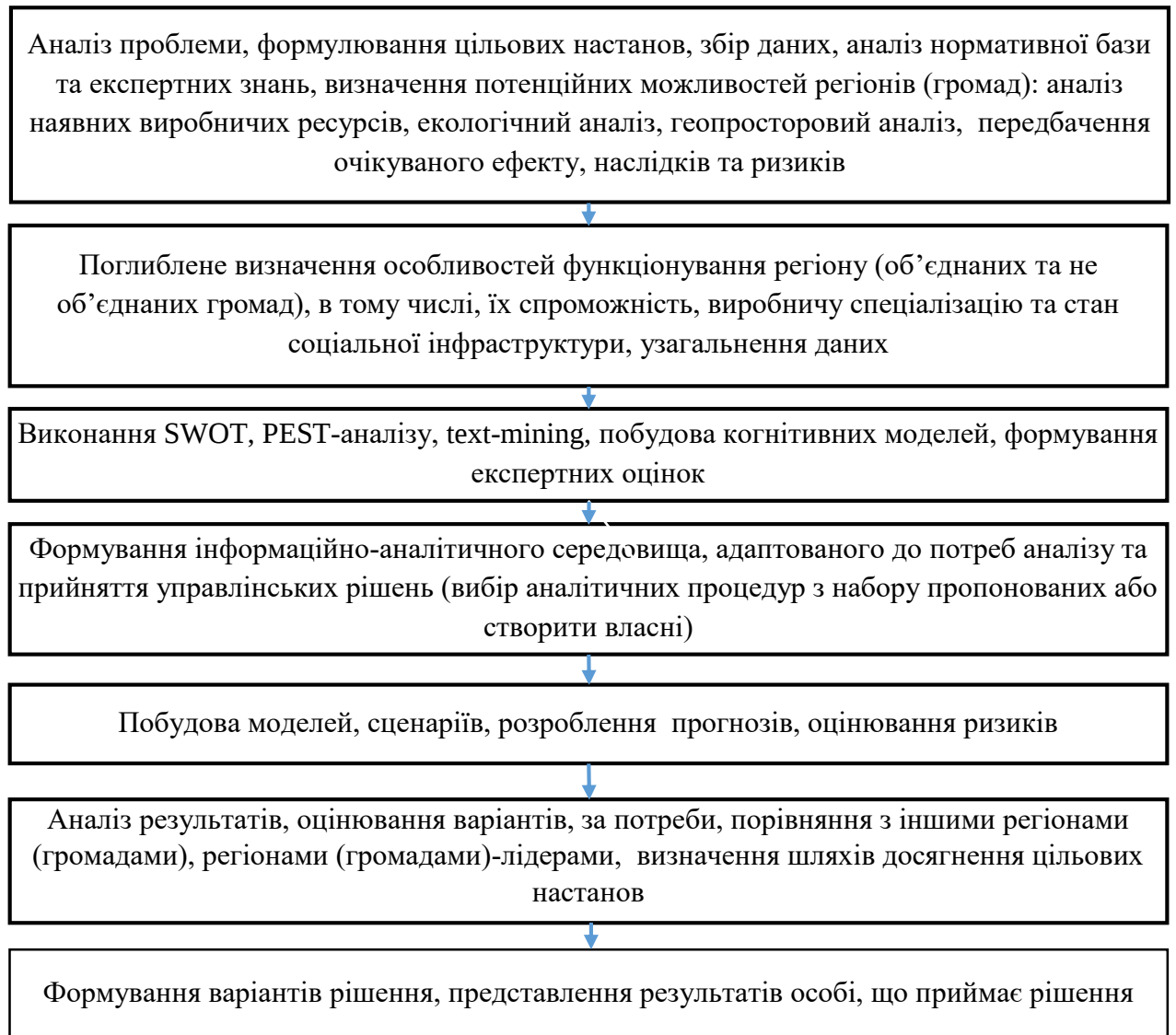


Рисунок 2.3 - Загальна (спрощена) схема застосування інформаційних технологій для підтримки прийняття рішень в управлінні регіональним розвитком

Розглядаючи процес прийняття рішень в управлінні регіональним розвитком, враховано, що інформаційне середовище системи є динамічним, його учасники - різні суб'єкти - органи державного та публічного управління, які мають різне підпорядкування, територіальні громади, бізнес, тощо, мають різні цілі, компетентність, ситуаційну обізнаність. Рішення приймаються децентралізовано, відповідно в межах повноважень рівня управління, відповідно до стратегічних, оперативних та тактичних цілей. Тому, для забезпечення децентралізації та інтеперабельності розроблюваних інформаційних технологій запропоновано використання мережецентричного підходу, який

забезпечить узгоджену взаємодію учасників прийняття рішень в управлінні регіональним розвитком.

Тому, для забезпечення сучасних вимог до інформаційних систем регіонального рівня пропонується застосування мережецентричної архітектури. Дана архітектура більше характерна для військової сфери, однак, для вирішення задач підтримки прийняття рішень в управлінні регіональним розвитком вона має низку переваг. Це й гнучка архітектура, яка дозволяє нарощувати систему, доповнюючи та видаляючи окремі ланки. З точки зору організації управлінського процесу, ієрархія управління «центр-регіон-район-громада» та взаємодія об'єктів управління, наприклад «громада-громада», «громада-регіон», «громада-район» здатна до самоорганізації, може динамічно змінюватись, адаптуватись під конкретну задачу. Крім того, обмін інформацією між елементами системи здійснюється не лише за підпорядкуванням, а й між всіма елементами системи, залежно від того, яка саме інформація потрібна, який ресурс треба задіяти для розв'язання конкретної управлінської задачі. Схема інформаційної взаємодії учасників системи управління регіональним розвитком представлена на рис. 2.4.

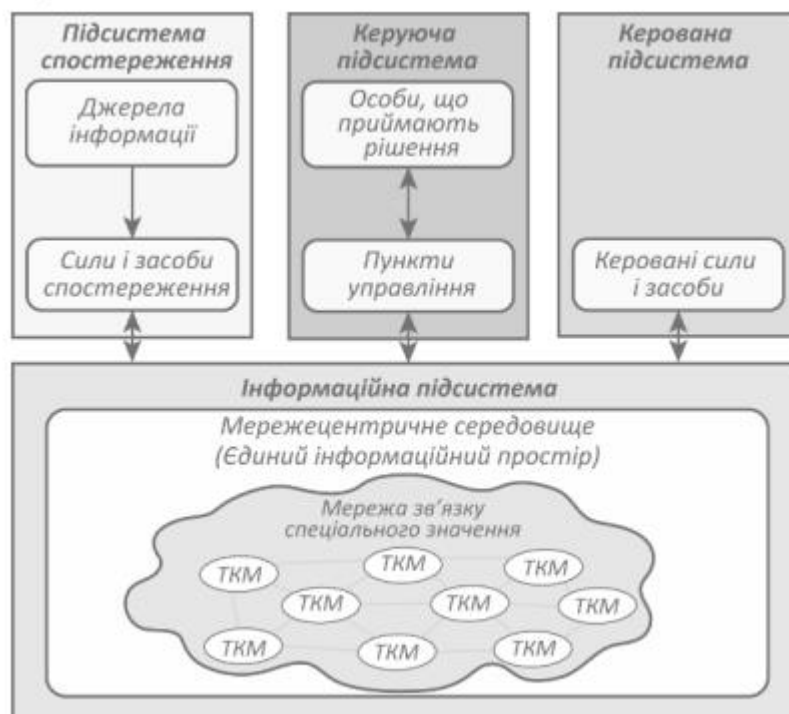


Рисунок 2.4 – Узагальнена схема мережецентричної системи управління []

Мережецентричний підхід дозволяє врахувати й таку особливість регіонального рівня, як організація управління у територіальних громадах та їх взаємодія з районним та регіональним рівнями управління. Громади приймають власні рішення стосовно стратегій розвитку та реалізації власних повноважень, кооперативні стратегії об'єктів рівня громади – бізнесу, місцевого самоврядування, держави, суспільства тому на рівні регіону, управління регіональним розвитком зосереджується більше на координації процесів, узгодженні кооперативних стратегій об'єктів «громада-район-регіон» загальної стратегії (програми).

Задача розроблення інформаційної технології управління регіональним розвитком розглядається як з точки зору забезпечення реалізації якісної аналітичної підсистеми, так і мінімізації витрат на забезпечення реалізації та впровадження, можливості забезпечення роботи з її складовими для посадовців, які не мають навичок аналітичної діяльності.

Специфічними аспектами пропонованої системи є необхідність забезпечення узгодженої взаємодії органів державного управління, місцевого самоврядування, підприємств, установ, організацій, розташованих на певній території, забезпечення моніторингу всіх етапів циклу управління регіональним розвитком від формування планів та стратегій до аналізу результативності системи управління регіоном та підтримки прийняття рішень у кризових ситуаціях

Система має функціонувати використовуючи наявні організаційні структури регіонального рівня та рівня територіальних громад, мати уніфіковану реалізацію на місцях, єдине методичне, правове, апаратно-технічне та програмне забезпечення. Система має використовувати єдиний інформаційний простір, користувачі якого мають доступ до баз даних та знань використовуючи єдину комп'ютерну мережу. Комп'ютерна мережа повинна забезпечувати спільне використання ресурсів та хмарних сервісів, а також захист даних, які не перебувають у вільному доступі. Система має функціонувати як геопросторова

розподілена інформаційна система, бути сумісною з інформаційно-аналітичними системами міністерств та органів державного управління відповідного рівня.

Такий підхід дозволить оперативно та якісно опрацювати великі обсяги даних щодо соціально-економічного розвитку регіонів та громад, фінансового забезпечення місцевого самоврядування для моніторингу соціально-економічного розвитку територій, дослідження досвіду різних регіонів та громад, тощо для координації взаємодії «громади-регіони-центр», надання цільової державної підтримки, адаптованої до потреб конкретних громад та регіонів, народження ефективного державно-приватного партнерства, удосконалення планової та аналітичної роботи органів державного управління і місцевого самоврядування.

2.3 Методологія застосування мережцентричного підходу для побудови інформаційних технологій управління регіональним розвитком

Формування нової моделі регіонального управління, яка передбачає зміну управлінської структури регіону від жорстко ієрархічної до більш гнучкої, коли територіальні громади, бізнес, громадськість отримують можливість брати активну участь у регіональних процесах, потребує нового підходу до організації підтримки прийняття рішень. Виникає потреба у створенні адаптивних інтелектуальних систем підтримки прийняття рішень, які можуть бути налаштовані до умов певної предметної області та вимоги осіб, що приймають рішення. На рівні регіону та громади вже впроваджено значну кількість інформаційних систем, спрямованих на вирішення задач надання адміністративних послуг населенню та автоматизацію документообігу в органах державного управління та місцевого самоврядування. Крім того, вирішуючи задачі інформатизації, більшість територіальних громад, органів управління районного та обласного рівнів мають власні напрацювання щодо автоматизації процесів підтримки прийняття рішень, планування та аналізу результатів виконання бюджетів, програм та стратегій розвитку. Вузькоспеціалізовані

інформаційні системи не знайшли підтримки на місцях, оскільки вони не здатні швидко адаптовуватись до умов реформування системи управління та розширення кола управлінських задач. Крім того, програмні рішення, які не можливо швидко оновлювати не задовольняють користувачів із-за неможливості їх використання у період оновлення. Сьогодні потрібні гнучкі адаптивні рішення, які можуть працювати з інформацією, отриманою з різних джерел, використовуючи її для аналізу, побудови моделей та прогнозів, крім того, важливим питанням є створення інтерфейсу користувача, який можливо швидко налаштувати під різні задачі, вирішувані у системі підтримки прийняття рішень. Для регіонального рівня управління важливим питанням є узгодження та координування співпраці з територіальними громадами, районами, бізнесом, громадськістю. За таких умов актуальним є застосування мережецентричного підходу до управління інформаційним забезпечення регіонального розвитку. Цей підхід відповідає в значній мірі сучасній концепції управління регіональним розвитком, оскільки передбачає реалізацію як ієрархічного, так і мережецентричного керування. Осередком такої системи управління є органи місцевого самоврядування та державного управління регіону, а територіальні громади є окремими центрами управління із своїми цілями й задачами. Районний рівень виконує скоріше координуючу функцію, яка згодом, в процесі реформування місцевого самоврядування до норм європейського співтовариства може стати не потрібною. Перевагами застосування мережецентричного підходу є те, що його основу становить центральна ланка – органи управління регіоном, система являє собою багат шарову структуру – «регіон-район-громада». горизонтальні зв'язки кожного рівня створюють замкнені системи з власними ресурсами, структурою, зв'язками. Використання такої концепції побудови інформаційних технологій для системи підтримки прийняття рішень в управлінні регіональним розвитком дозволяє перейти від багаторівневої ієрархічної моделі управління до розподіленої мережевої системи. Така система дозволить зменшити час на проходження ієрархії для узгодження прийняття рішень, що в умовах військового стану є особливо актуальним. Застосування

мережецентричного підходу потребує формування єдиного інформаційного простору. Крім того, сучасний інформаційний простір використовує різні джерела даних, дані у яких можуть мати різні типи представлення даних – бути структурованими та не структурованими. Тому в системі важливо забезпечити реалізацію адаптивної бази даних, яка може використовуватись, навіть коли система буде нарощуватись.

Отже, управління за мережецентричного підходу передбачає наявність концентричної структури, центром якої є кореневий вузол системи, що вибудовується відповідно до інформаційної моделі системи. Особливістю мережецентричного підходу є організація взаємодія між об'єктами у інформаційній структурі системи за допомогою механізму посилення об'єктів. На наступному рівні концентричної структури визначаються типи об'єктів системи, з множини типів, визначених у інформаційній системі, вони поєднуються зв'язками, таким чином, що у об'єкта одна з властивостей може бути посиленням і вказувати на об'єкт заданого типу. Таким чином, використовуючи систему посилення між об'єктами встановлюються зв'язки «напрямую», без урахування ієрархії, що дозволяє вільно переходити від ієрархічної системи до мережецентричної. Таким чином, відбувається формування єдиного інформаційного простору, до складу якого входять джерела інформації, елементи системи, дані, що зберігаються.

Перевагами такої організації системи те, що алгоритми функціонування системи можуть бути адаптовані до змін у зовнішньому середовищі системи або у відповідь на дії користувача. У системі управління регіональним розвитком працює значна кількість користувачів, всі мають різні задачі, цільові настанови, дані, використовувані у роботі. Тому запропонована структура системи передбачає налаштування на певну предметну область, використання даних за потреб користувача з різних джерел та предметних областей. Масштабованість мережецентричної системи забезпечується як вертикально, так і горизонтально. Тобто, на рівні регіону має функціонувати потужний даних центр, який забезпечує обмін інформаційними потоками в мережі в режимі реального часу. Крім того,

мережецентричний підхід до організації інформаційних систем за рахунок розподіленої обробки даних, забезпечення динамічної маршрутизації передавання даних у комп'ютерній мережі дозволяє не лише масштабувати обчислювальні ресурси, а й забезпечувати стійкість системи до збоїв за рахунок можливості використання резервних каналів зв'язку. При побудові такої системи вирішується проблема інтеоперабельності взаємодії складових системи - адже складові функціонують на основі єдиної методики, використовуючи єдині стандарти.

Цілеспрямована взаємодія складних систем нижчого рівня - територіальних громад здійснюється на основі принципів самоорганізації та створює синергетичний ефект від взаємодії. Складові системи автономні – громади, райони, регіони, функціонують в єдиному інформаційному просторі, мають свої цілі, пріоритети, можуть існувати за рахунок власних ресурсів, встановлювати як вертикальні так і горизонтальні зв'язки. Крім того, мережецентрична модель відтворює фрактальність регіональних соціально-економічних систем, що створює передумови для розроблення типових програмно-апаратних рішень, що об'єднані організаційно та технічно на основі єдиного інформаційного середовища. Основою для розроблення інформаційних технологій є державні стандарти України, технічні вимоги та методи забезпечення інтеоперабельності і сумісності наборів геопросторових даних та геоінформаційних сервісів [5], документи консорціуму NCOIC щодо проектування мережецентричних інформаційно-управляючих систем, зокрема модель оцінки інтеоперабельності системи SCOPE-модель, керівництво з забезпечення мінімального рівня інтеоперабельності MLI та керівництво по забезпеченню інтеоперабельності NIF [24]. Важливою складовою розроблюваної інформаційної технології є комплексна система захисту інформації, яка включає програмно-апаратні та програмні засоби технічного та криптографічного захисту інформації, необхідні для забезпечення належного рівня захищеності інформаційних ресурсів системи, побудована згідно законодавчо встановлених вимог та має атестат відповідності .

Оскільки, значна увага приділяється роботі з даними, то пропонується для розв'язання цієї задачі пропонується застосувати конвергентний підхід до організації збору, оброблення та зберігання даних. Тобто, має бути сформована система, центральною частиною якої є центр оброблення даних, а складовими - система обміну даними між компонентами, хмарні сервіси. Дані для аналізу використовуватимуться як з аналітичного сховища, так і в результаті виконання API-запитів, в тому числі, коли потрібно здійснювати обробку даних у режимі реального часу.

Користувачі системи мають бути розподілені за групами, які мають різні задачі, повноваження та права доступу. Для забезпечення надійності віддаленого доступу, взаємодії між користувачами різних рівнів, надання їм доступу відповідно до повноважень та прав, рівня системи кібербезпеки, система повинна мати підсистеми та адміністрування.

Використання пропонованої системи забезпечить скорочення часу на оброблення значних обсягів даних (структурованих та неструктурованих) та дозволить отримувати результати аналізу та прогнозування високої якості у режимі реального часу, що покращить якість управлінських рішень як на рівні регіону, так і громади. Крім того, використання єдиного рішення дозволить скоротити витрати на реалізацію програм інформатизації у регіонах, посилить зв'язки між регіонами та громадами, дозволить покращити аналітичну складову місцевого та регіонального управління.

2.4 Моделі та методи виявлення загроз та реагування на ризики загрози в інформаційній технології прийняття рішень

Сучасне IT-середовище об'єктів регіонального та місцевого рівнів, а також систем обробки державних інформаційних ресурсів, являє собою складну систему, що об'єднує різноманітні інформаційні, нормативні, програмні, технічні, людські й інші види ресурсів для забезпечення досягнення цілей

сталого розвитку. Це зумовлює зростання потреби у підвищенні ефективності використання захищених ІТ та інформаційних систем різних класів, збільшення переваг і усунення недоліків від їх застосування, а також встановлює критерії обґрунтування витрат на забезпечення захищених інформаційно-комунікаційних систем (ІКС; нормативно-правове, організаційне, технічне забезпечення).

Перший напрямок, пов'язаний із впровадженням аудиту, супроводження сучасних інформаційних систем для державного та регіонального управління та їх безпеки. Організація системи аудиту і контролю є ключовим компонентом для забезпечення якості функціонування захищених інформаційних систем регіону. Без надійних комплексних систем захисту інформації та результативних заходів системи аудиту і контролю, вона не в змозгу надійно і якісно виконувати базові процеси надання інформаційних послуг, забезпечувати процеси захисту інформаційних ресурсів різних класів, виконувати операції/транзакції та узагальнювати надійну діяльність і звітність тощо [1-4, 72-75]. При розв'язуванні задач, що пов'язані з забезпеченням кібербезпеки та взаємодії складових розгалужених комп'ютерних мереж інформаційної системи, виділяють декілька напрямків.

Інший напрям пов'язаний з проблемами протидії тероризму та актам незаконного втручання у діяльність об'єктів системи державного та регіонального управління, територіальних громад. Класифікація моделей протидії тероризму знайшла своє відображення у ряді робіт [40,76], присвячених проблемі моделювання системи протидії тероризму. До концептуальних моделей пропонується [40] віднести моделі, які розробляються фахівцями предметної області, політологами, психологами, соціологами [76]. В роботі [77] подані емпіричні дані щодо моделей прийняття рішень учасниками терористичних організацій на різних рівнях: стратегічному, тактичному та оперативному, а також на рівні окремого терориста.

Моделі аналізу і синтезу, як правило, є математичними чи фізичними моделями. В огляді [77-78] охарактеризовано і наведено більшість таких робіт. Зокрема у роботі [78] розроблена основа для розвитку класифікації

терористичних угруповань хімічної, біологічної, радіологічної та ядерної зброї з використанням евристичного методу розпізнавання образів, методу дерев класифікації і дискримінаційного аналізу. Стосовно систем безпеки на транспорті ряд робіт присвячений аналізу пристроїв задля підвищення імовірності виявлення і зниження інтенсивності помилкових тривог. У роботі загроз і визначення пріоритетів заходам безпеки об'єктам.

Складність реальних ситуацій, пов'язаних із забезпеченням безпеки, потребують універсальності застосованих математичних моделей.

Ці вимоги неминуче приводять у протиріччя із надійністю та обґрунтованістю результатів моделювання, тому при реалізації моделей у вигляді ієрархії (більш низьким рівням ієрархії відповідає більш висока ступінь деталізації опису модельованих систем) або горизонтальному ланцюжку, у кожному елементі якого ступінь деталізації приблизно однаковий.

В роботах [] детально розглянуті рівня моделювання (ієрархії моделей) протидії кібертероризму. Теоретико-ігрові моделі боротьби з тероризмом подані у роботах [80-82].

В роботі [83] розглянуті можливості застосування моделей і методів дослідження операцій для планування захисту об'єктів інформатизації. Адаптація цих моделей включає врахування стохастичної, інформаційної і поведінкової невизначеностей терористів. У цій роботі [78], зокрема, узагальнено задачі антагоністичної гри нападу і захисту та оптимального розподілу захисних ресурсів.

Прикладом використання комплексних моделей з параметрами, вимірюваними у різних шкалах, є теоретико-ігрова модель ARMOR [84]. У роботі [] представлений опис тестів для перевірки ARMOR, який експлуатується ще з 2007 р.

Системи масового обслуговування (СМО) знайшли своє застосування у математичних моделях інформаційних технологій комп'ютерних мереж. Параметр η є інтенсивністю покидання заявок із каналу обслуговування СМО

внаслідок обмеження їхнього перебування в системі. Параметр μ характеризує систему протидії $\mu = \frac{1}{t_{ox}}$ [зловм/од. часу], де t_{ox} - середній час застосування охоронником засобів протидії зловмиснику. Протидія зловмисникам з боку системи захисту може бути організована різними способами, що обумовлює тип системи масового обслуговування. До першої групи СМО належать:

1. СМО М/М/п/т з обмеженням ($\eta \neq 0$), без взаємодопомоги ($h = n, g = 1$), неординарними заявками і випадковим числом заявок у групі. Тут h величина, що дорівнює відношенню всієї кількості n каналів обслуговування до кількості g елементів системи захисту, які об'єднані у групу для одному зловмиснику, тобто

2. СМО М/М/п/т з обмеженням ($\eta \neq 0$), повною взаємодопомогою ($h = 1; g = n$), неординарними заявками і випадковим числом заявок у групі.

3. СМО М/М/п/т з обмеженням ($\eta \neq 0$), з частковою взаємодопомогою (неординарними заявками і випадковим числом заявок у групі).

До другої групи належать СМО немарковського типу, які моделюють умови, коли всі засоби захисту не знаходяться на об'єкті, наприклад, коли необхідно зосередити додаткові засоби. Тобто, процес протидії складається з двох фаз тривалістю відповідно t_1 - часу зосередження і t_{ox} - часу застосування засобів протидії, де t_1 має показниковий розподіл з параметром $\mu_1 = \frac{1}{t_1}$ [охор./од. часу], а t_2 - з параметром $\mu_2 = \frac{1}{t_{ox}}$ [зловм./од. часу]. Тобто сумарний час протидії має узагальнений розподіл Ерланга з параметрами μ_1 і μ_2 .

Такі СМО мають обмеження ($\eta \neq 0$), можуть бути з різними характеристиками взаємодопомоги, є система масового обслуговування з неоднорідними заявками і випадковим числом заявок у групі.

Основні аспекти математичних моделей цих СМО розглянуті у роботах [86- інформаційної-комунікаційної мережі регіону пропонується до застосування у складі існуючої системи, регламентованою державними стандартами та нормативними документами державних органів у сфері захисту інформації.

Для оцінювання надійності функціонування розгалуженої комп'ютерної мережі регіону пропонується застосування математичного апарату процесів Маркова – моделювання процесу народження і загибелі, що може відбуватися як із дискретними, так і з неперервними змінами часу t [89]. При цьому головною умовою, яка неодмінно має виконуватися, що переходи процесу можливі лише до сусідніх станів.

Марковський процес у цьому разі описує зміни, які відбуваються в часі в певному обсязі популяції, а саме — зміни кількості одиниць певного виду складових популяції.

Розглянемо вид процесу народження та загибелі (або розмноження та вимирання) для задач функціонування об'єктів комп'ютерної мережі мережецентричної системи управління регіональним розвитком. Нехай, система S складається з n вузлів. В якості вузлів можна розглядати окремі елементи комп'ютерної мережі, що пов'язує об'єкти мережецентричної системи управління регіональним розвитком. Кожен із вузлів може незалежно від інших виходити з ладу, тобто на кожен із вузлів діє найпростіший “потік відмов”, подіями якого є відмови вузла. У проміжку між двома сусідніми відмовами, вузол працює безвідмовно. Середній час безвідмовної роботи кожного з вузлів позначимо через T . Якщо, в початковий момент часу, вузол справний, то після появи першої події потоку відмов, що діє на вказаний вузол, він виходить з ладу. При розгляді цього виду процесу «народження та загибелі» розглядають “весь” потік відмов, що дозволяє говорити про інтенсивність потоку відмов.

В якості потоку відмов можна розглядати як потік відмов, так і потік атак на об'єкти критичної інфраструктури.

Вузол, який вийшов з ладу, починає відразу відновлюватись. Вважаємо, що на кожний вузол діє найпростіший потік “відновлення”, подіями якого є відновлення роботи вузла, тобто закінчення ремонту вузла. На інтервалі між двома сусідніми відновленнями, вузол знаходиться в стані ремонту. Середній час відновлення (ремонту) позначаємо як T_v . Вузол, що знаходиться в ремонті, відновлюється після появи першої події потоку відновлень, що діє на вказаний

вузол. Як і у випадку потоку відмов, розглядають “весь” потік відновлень, що дозволяє стверджувати про інтенсивність даного потоку.

Припускається, що процес народження певної кількості одиниць організмів може бути змодельований пуассонівським процесом (поток) із параметром λ_k , що є інтенсивністю народження k -го організму, а загибель — експоненціальним законом розподілу з параметром λ_{μ_k} , де μ_k — інтенсивність загибелі k -го організму. Такі припущення добре узгоджуються з реальними процесами «народження і загибелі», які відбуваються в певному обмеженому просторі популяції, а також дають змогу застосовувати математичний апарат, будуючи такі ймовірнісні моделі, котрі можна використати для розв’язання широкого кола задач, що постають у системах обслуговування.

Для процесу «народження та загибелі», як уже відзначено, можливі лише переходи зі стану ω_k до стану ω_{k-1} або ω_{k+1} .

Нехай, обсяг популяції — k одиниць, а отже, процес перебуває у стані ω_k . Перехід процесу зі стану ω_k до стану ω_{k+1} відповідає випадковій події — народження з певною ймовірністю, однієї одиниці організму, а перехід процесу зі стану ω_k до стану ω_{k-1} відповідатиме випадковій події: загибель із певною ймовірністю однієї одиниці організму.

Три можливі переходи процесу розмноження та загибелі, де відсутні переходи процесу до несусідніх станів, представлені на рис. 2.5.

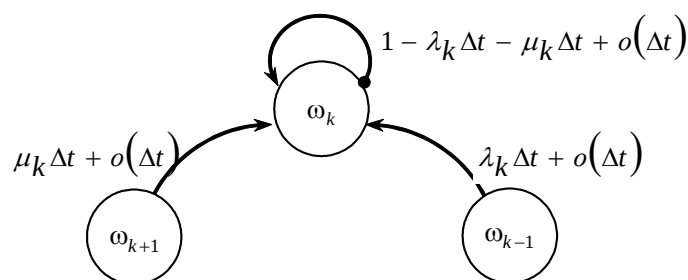


Рисунок 2.5 – Граф процесу розмноження і загибелі [86]

Для побудови ймовірнісної моделі цього процесу, доільно докладніше зупинитись на його переходах до сусідніх станів і відповідних імовірностях переходів.

Нехай у момент часу t було k одиниць популяції, тобто процес розмноження та загибелі перебував у стані ω_k .

Тоді, в момент часу $t + \Delta t$, де Δt — достатньо мала величина, процес перебуватиме у стані ω_k , якщо відбудеться одна з несумісних випадкових подій із відповідними ймовірностями, а саме:

- якщо в момент часу t обсяг популяції дорівнював $k + 1$ одиницям, то протягом часу Δt одна одиниця неодмінно має загинути з імовірністю

$$p_{k+1}(\Delta t) = \mu_{k+1}\Delta t + a(\Delta t); \quad (2.1)$$

- якщо в момент часу t обсяг популяції дорівнював $k - 1$ одиниць, то протягом часу Δt одна одиниця неодмінно народиться з імовірністю

$$p_{k-1}(\Delta t) = \lambda_{k-1}\Delta t + a(\Delta t); \quad (2.2)$$

- якщо в момент часу t обсяг популяції дорівнював k одиниць і протягом часу Δt ця кількість не зміниться, тобто жодна одиниця не загинула і не народилась із імовірністю

$$p_k(t + \Delta t) = 1 - \lambda_k\Delta t - \mu_k\Delta t + a(\Delta t). \quad (2.3)$$

Наведені переходи та відповідні їм імовірності зображено на рис. 2.6 для $k \geq 1$.

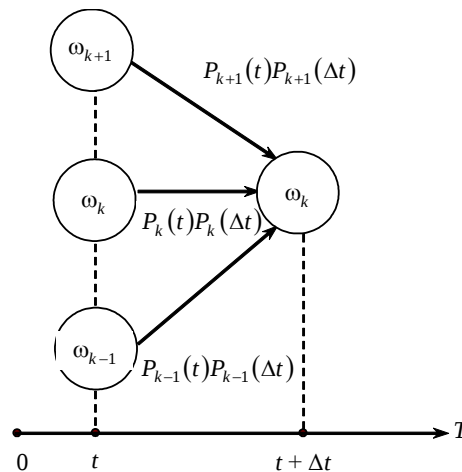


Рисунок 2.6 – Переходи процесу розмноження і загибелі

Для $k = 0$ переходи процесу та відповідні їм імовірності зображено на рис. 2.7.

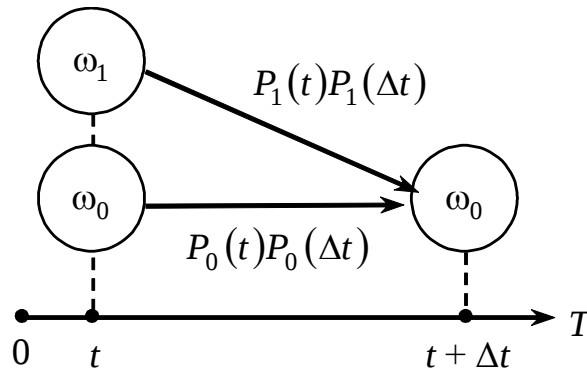


Рисунок 2.7 – Ймовірності переходів процесу розмноження і загибелі

На основі наведених міркувань отримаємо систему рівнянь (2.4):

$$\begin{cases} p_k(t + \Delta t) = p_k(t)p_k(\Delta t) + p_{k+1}(t)p_{k+1}(\Delta t) + p_{k-1}(t)p_{k-1}(\Delta t), & k \geq 1, \\ p_0(t + \Delta t) = p_0(t)p_0(\Delta t) + p_1(t)p_1(\Delta t). \end{cases} \quad (2.4)$$

При цьому

$$\sum_{k=0}^{\infty} p_k(t) = 1. \quad (2.5)$$

Скориставшись (2.1)—(2.3), систему (2.4) буде подана у вигляді

$$\begin{cases} p_0(t + \Delta t) = p_0(t)(1 - \lambda_0\Delta t + a(\Delta t)) + p_1(t)(\mu_1\Delta t + a(\Delta t)), \\ p_k(t + \Delta t) = p_k(t)(1 - \lambda_k\Delta t - \mu_k\Delta t + a(\Delta t)) + \\ + p_{k+1}(t)(\mu_{k+1}\Delta t + a(\Delta t)) + p_{k-1}(t)(\lambda_{k-1}\Delta t + a(\Delta t)), & k \geq 1. \end{cases} \quad (2.6)$$

Розкривши дужки в системі (2.6), маємо:

$$\begin{cases} p_0(t + \Delta t) = p_0(t) - \lambda_0\Delta tp_0(t) + p_0(t)a(\Delta t) + \mu_1\Delta tp_1(t) + p_1(t)a(\Delta t), \\ p_k(t + \Delta t) = p_k(t) - \lambda_k\Delta tp_k(t) - \mu_k\Delta tp_k(t) + p_k(t)a(\Delta t) + \mu_{k+1}\Delta tp_{k+1}(t) + \\ + p_{k+1}(t)a(\Delta t) + \lambda_{k-1}\Delta tp_{k-1}(t) + p_{k-1}(t)a(\Delta t), \end{cases}$$

або

$$\begin{cases} p_0(t + \Delta t) = p_0(t) - \lambda_0\Delta tp_0(t) + \mu_1\Delta tp_1(t) + (p_0(t) + p_1(t))a(\Delta t); \\ p_k(t + \Delta t) = p_k(t) - (\lambda_k + \mu_k)\Delta tp_k(t) + \mu_{k+1}\Delta tp_{k+1}(t) + \\ + \lambda_{k-1}\Delta tp_{k-1}(t) + (p_k(t) + p_{k+1}(t) + p_{k-1}(t))a(\Delta t). \end{cases} \quad (2.8)$$

Від системи (2.9) переходимо до системи

$$\begin{cases} p_0(t + \Delta t) - p_0(t) = -\lambda_0 \Delta t p_0(t) + \mu_1 \Delta t p_1(t) + (p_0(t) + p_1(t))a(\Delta t), \\ p_k(t + \Delta t) - p_k(t) = -(\lambda_k + \mu_k) \Delta t p_k(t) + \mu_{k+1} \Delta t p_{k+1}(t) + \\ + \lambda_{k-1} \Delta t p_{k-1}(t) + (p_k(t) + p_{k+1}(t) + p_{k-1}(t))a(\Delta t). \end{cases} \quad (2.9)$$

Поділивши ліву і праву частини системи на Δt , дістанемо:

$$\begin{cases} \frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} = -\lambda_0 p_0(t) + \mu_1 p_1(t) + (p_0(t) + p_1(t)) \frac{a(\Delta t)}{\Delta t}, \\ \frac{p_k(t + \Delta t) - p_k(t)}{\Delta t} = -(\lambda_k + \mu_k) p_k(t) + \mu_{k+1} p_{k+1}(t) + \lambda_{k-1} p_{k-1}(t) + \\ + (p_k(t) + p_{k+1}(t) + p_{k-1}(t)) \frac{a(\Delta t)}{\Delta t}, \quad k \geq 1. \end{cases}$$

Переходячи до границі за $\Delta t \rightarrow 0$, отримаємо:

$$\begin{cases} \frac{dp_0(t)}{dt} = -\lambda_0 p_0(t) + \mu_1 p_1(t), \\ \frac{dp_k(t)}{dt} = -(\lambda_k + \mu_k) p_k(t) + \mu_{k+1} p_{k+1}(t) + \lambda_{k-1} p_{k-1}(t), \quad k \geq 1, \end{cases}$$

оскільки

$$\lim_{\Delta t \rightarrow 0} \frac{a(\Delta t)}{\Delta t} = 0,$$

або

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t), \\ p'_k(t) = -(\lambda_k + \mu_k) p_k(t) + \mu_{k+1} p_{k+1}(t) + \lambda_{k-1} p_{k-1}(t), \quad k \geq 1. \end{cases} \quad (2.10)$$

Отже, отримано систему диференціально-різницьких рівнянь (2.10), яка описує динаміку ймовірнісного процесу «народження і загибелі».

Зосередивши увагу, наприклад, на стані ω_k , побачимо, що перейти до цього стану можна лише зі станів ω_{k-1} або ω_{k+1} , і навпаки, залишаючи стан ω_k , можна потрапити лише в стан ω_{k-1} або ω_{k+1} . Тобто, таким чином можна передбачити, наскільки надійною буде розроблена система

Серед процесів, важливих для функціонування комп'ютерних систем та мереж особливе місце належить проблемі опрацюванню черг. Таке явище, як черга, певною мірою є випадковим, причому виникає воно практично в усіх сферах життя та діяльності людини. Вироби, які чекають на обслуговування на

конвеєрах, літаки — на звільнення злітної смуги аеродрому, судна — на розвантаження в морських або річкових портах, пакети у вузлах комп'ютерної мережі також утворюють черги. Будь-яка система, яка пов'язана з певною операцією обслуговування, завжди має справу з таким явищем, як черга.

Де б не виникали черги, актуальною є задача: якою має бути максимально допустима довжина черги, з якою швидкістю мають обслуговуватися клієнти, тощо. Вони мають узагальнюючу назву — вимоги [89]. Такі питання: як, скільки потрібно обслуговуючих пристроїв, стосуються саме каналу обслуговування.

Для розв'язання таких задач, потрібно побудувати таку математичну модель, за допомогою якої можна відтворити всі можливі ситуації, пов'язані з функціонуванням системи, що досліджуються. Для цього необхідно провести серію експериментів, отримати результати моніторингу, а використовуючи відповідний математичний апарат, змоделювати ситуацію. І тоді, використовуючи дані, здобуті в результаті спостережень за функціонуванням реальних систем, можна перевірити, чи адекватними є створені математичні моделі. Вперше такі моделі побудував А. К. Ерланг [90]. У його моделі було введено характеристики типових черг. Результати було узагальнено в теорію, яку згодом розвивали багато інших учених.

З огляду на те, що процес утворення черг і час, який витрачається каналом для обслуговування, мають випадковий характер, моделі, створені для дослідження систем, є ймовірнісними, або стохастичними, а самі системи називають системами масового обслуговування (СМО).

Основні компоненти СМО такі [89]:

- вхідний потік вимог — вимоги надходять на обслуговування у випадкові моменти часу і утворюють пуассонівський потік із інтенсивністю λ ;
- механізм обслуговування — час обслуговування є випадковою величиною, що має експоненціальний закон розподілу із параметром μ , який розглядається як інтенсивність обслуговування вимог [89].

При цьому, якщо $\rho = \frac{\lambda}{\mu} > 1$, тобто завантаження системи більше за одиницю, то черга має тенденцію постійно збільшуватися і система не зможе

набути стаціонарного стану. За $\rho < 1$ черга має тенденцію до зменшення і система при цьому зможе досягти стаціонарного стану (режиму) й тоді існують стаціонарні ймовірності.

- дисципліна обслуговування вимог — дисципліна обслуговування черги встановлюється в залежності від специфіки роботи системи. Найпростіша й логічніша — перший прийшов і першим обслужився, а за наявності черги — чекаєш свого часу на обслуговування. В інших випадках можуть бути введені пріоритети в обслуговуванні вимог. У цьому випадку вимоги поділяють на пріоритетні і прості. В окремих випадках, можуть бути, введені кілька рівнів пріоритетності [89].

Так, наприклад, за двох рівнів пріоритетності вимоги поділяють на прості, з відносним пріоритетом, із абсолютним пріоритетом. Вимоги із відносним пріоритетом мають перевагу в обслуговуванні над простими. А вимоги з абсолютним пріоритетом, мають перевагу в обслуговуванні як над простими вимогами, так і над вимогами із відносним пріоритетом [89].

Отже, можна зробити висновок: між процесами, що відбуваються за певного рівня формалізації в системах масового обслуговування, і процесами, які фіксують зміни, що відбуваються в певному обсязі популяції, за умови $\mu_k = \mu$, $\lambda_k = \lambda$, існує аналогія. І дійсно, якщо в обсязі популяції зміна на одиницю моделюється пуассонівським потоком (обсяг популяції збільшується на одиницю) або експоненціальним законом (обсяг популяції зменшується на одиницю), то в системах масового обслуговування зміна на одиницю вимог в системі також буде моделюватися пуассонівським потоком (одна вимога надійшла в систему) або експоненціальним законом розподілу (одна вимога обслужилась і покинула систему).

Отже, ймовірнісна модель (3.20), яка створювалася для процесу народження і загибелі певного обсягу популяції, повністю відповідає процесу, що має місце в системі масового обслуговування з одним обслуговуючим каналом, одним пуассонівським потоком вимог та експоненційним законом обслуговування цих

вимог, і є фундаментом, на основі якого створюються ймовірнісні моделі складніших систем обслуговування.

Якщо в системі (2.11) визначити $\mu_k = 0$, для всіх значень k , то маємо процес «чистого» розмноження:

$$\begin{cases} \frac{dp_k(t)}{dt} = -\lambda_k p_k(t) + \lambda_{k-1} p_{k-1}(t), & k \geq 1, \\ \frac{dp_0(t)}{dt} = \lambda_0 p_0(t). \end{cases} \quad (2.11)$$

Якщо в системі (2.12) $\mu_k = \mu = \text{const}$ і $\lambda_k = \lambda = \text{const}$, то вона набуде такого вигляду:

$$\begin{cases} \frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t), \\ \frac{dp_k(t)}{dt} = -(\lambda + \mu) p_k(t) + \lambda p_{k-1}(t) + \mu p_{k+1}(t), & k \geq 1. \end{cases} \quad (2.12)$$

Система (2.12) - моделю Ерланга. Незалежний від часу, її розв'язок можна отримати двома шляхами: розв'язати систему рівнянь (2.12), а потім у знайдених розв'язках спрямувати $t \rightarrow \infty$; 2) прирівняти до нуля похідні в системі (2.12) і здобути при цьому систему алгебраїчних рівнянь розв'язати відносно ймовірностей.

Знайдені розв'язки відповідають стаціонарному стану процесу. Імовірності, що відповідають такому стану, називають стаціонарними, і їх значення не залежать від часу t :

$$p_0(t) \Big|_{t \rightarrow \infty} = p_0(\infty) = p_0, \quad p_k(t) \Big|_{t \rightarrow \infty} = p_k(\infty) = p_k.$$

Отже, у стаціонарному режимі (стані) система (3.22) набуде такого вигляду:

$$\begin{cases} 0 = -\lambda p_0 + \mu p_1, \\ 0 = -(\lambda + \mu) p_k + \mu p_{k+1} + \lambda p_{k-1}, & k \geq 1, \end{cases}$$

або

$$\begin{cases} \lambda p_0 = \mu p_1, \\ (\lambda + \mu)p_k = \mu p_{k+1} + \lambda p_{k-1}, \quad k \geq 1. \end{cases} \quad (2.13)$$

Візьмемо величину $\rho = \frac{\lambda}{\mu}$, яку називають завантаженням досліджуваного процесу. Стаціонарний стан процесу існуватиме лише за умови, що завантаження $\rho < 1$.

Оскільки $\frac{\lambda}{\mu} = \rho$, то система (2.14) набирає такого вигляду:

$$\begin{cases} p_1 = \rho p_0, \\ (1 + \rho)p_k = p_{k+1} + \rho p_{k-1}, \quad k \geq 1. \end{cases} \quad (2.14)$$

Послідовно розв'язуючи друге рівняння системи (2.14) для значень $k = 1, 2, 3, 4, \dots$, дістаємо:

$$\begin{aligned} k = 1: (1 + \rho)p_1 &= p_2 + \rho p_0 \rightarrow (1 + \rho)\rho p_0 = p_2 + \rho p_0 \rightarrow p_2 = \rho^2 p_0; \\ k = 2: (1 + \rho)p_1 &= p_3 + \rho p_1 \rightarrow (1 + \rho)\rho^2 p_0 = p_3 + \rho^2 p_0 \rightarrow p_3 = \rho^3 p_0; \\ k = 3: (1 + \rho)p_3 &= p_4 + \rho p_2 \rightarrow (1 + \rho)\rho^3 p_0 = p_4 + \rho^3 p_0 \rightarrow p_4 = \rho^4 p_0; \\ k = 4: (1 + \rho)p_4 &= p_5 + \rho p_3 \rightarrow (1 + \rho)\rho^4 p_0 = p_5 + \rho^4 p_0 \rightarrow p_5 = \rho^5 p_0; \\ &\dots\dots\dots \\ k = n: p_n &= \rho^n p_0. \end{aligned} \quad (2.15)$$

Скориставшись умовою нормування, дістанемо:

$$\sum_{k=0}^{\infty} p_k = 1 \rightarrow \sum_{k=0}^{\infty} \rho^k p_0 = 1,$$

звідки

$$p_0 \sum_{k=0}^{\infty} \rho^k = 1 \rightarrow p_0 = \frac{1}{\sum_{k=0}^{\infty} \rho^k}.$$

Оскільки

$$\sum_{k=0}^{\infty} \rho^k = 1 + \rho + \rho^2 + \rho^3 + \dots = \frac{1}{1-\rho},$$

то

$$p_0 = 1 - \rho, p_k = \rho^k (1 - \rho). \quad (2.16)$$

Визначимо основні числові характеристики для моделі Ерланга:

$$\begin{aligned} M(k) &= \sum_{k=0}^{\infty} k p_k = \sum_{k=0}^{\infty} k \rho^k (1 - \rho) = (1 - \rho) \sum_{k=0}^{\infty} k \rho^k = (1 - \rho) \rho \sum_{k=1}^{\infty} k \rho^{k-1} = \\ &= \left| \begin{array}{l} \text{Оскільки} \\ k \rho^{k-1} = (\rho^k)' \\ \text{то дістаємо} \end{array} \right| = (1 - \rho) \rho \left(\sum_{k=0}^{\infty} \rho^k \right)' = (1 - \rho) \rho \left(\frac{1}{1-\rho} \right)' = \frac{\rho(1-\rho)}{(1-\rho)^2} \\ &= \frac{\rho}{1-\rho}. \end{aligned}$$

Отже, математичне сподівання буде таке:

$$M(k) = \frac{\rho}{1-\rho}. \quad (2.17)$$

Щоб визначити дисперсію, знайдемо $M(k^2)$:

$$\begin{aligned} M(k^2) &= \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=0}^{\infty} k^2 \rho^k (1 - \rho) = (1 - \rho) \sum_{k=0}^{\infty} k^2 \rho^k = \\ &= (1 - \rho) \rho^2 \sum_{k=2}^{\infty} k(k-1) \rho^{k-2} + (1 - \rho) \rho \sum_{k=1}^{\infty} k \rho^{k-1} = \\ &= (1 - \rho) \rho^2 \left(\sum_{k=0}^{\infty} \rho^k \right)'' + (1 - \rho) \rho \left(\sum_{k=0}^{\infty} \rho^k \right)' = \\ &= \left| \begin{array}{l} \text{Оскільки} \quad k(k-1) \rho^{k-2} = (\rho^k)'' \\ \text{то} \quad k \rho^{k-1} = (\rho^k)' \end{array} \right| \\ &= (1 - \rho) \rho^2 \left(\frac{1}{1-\rho} \right)'' + (1 - \rho) \rho \left(\frac{1}{1-\rho} \right)' = \end{aligned}$$

$$= (1 - \rho) \rho^2 \frac{2}{(1 - \rho)^3} + (1 - \rho) \rho \frac{1}{(1 - \rho)^2} = \frac{2\rho^2}{(1 - \rho)^2} + \frac{1}{(1 - \rho)^2}.$$

Отже,

$$M(k^2) = \frac{2\rho^2}{(1 - \rho)^2} + \frac{\rho}{(1 - \rho)} = \frac{\rho^2 + \rho}{(1 - \rho)^2}.$$

Тоді дисперсія

$$D(k) = M(k^2) - M^2(k) = \frac{\rho^2 + \rho}{(1 - \rho)^2} - \frac{\rho^2}{(1 - \rho)^2} = \frac{\rho}{(1 - \rho)^2},$$

або

$$D(k) = \frac{\rho}{(1 - \rho)^2}. \quad (3.18)$$

Математичний апарат СМО часто використовується для вирішення питань моніторингу та реагування на загрози у комп'ютерних мережах різного призначення. Не виключення й система управління регіональним розвитком. Однак, для оцінювання ризиків і загроз в системі управління регіональним розвитком запропоновано використовувати представлений в роботі [110] інтегрований підхід до розробки скорингових карт. Використання даного підходу апробовано при розв'язанні задач оцінювання ризиків для об'єктів критичної інфраструктури в умовах зростагтя кількості кібератак та складності сучасних інформаційних систем. Підхід поєднує мультикритеріальний аналіз, який враховує технічні, економічні, екологічні та операційні показники, із технологіями штучного інтелекту та машинним навчанням.

Запропонована методологія передбачає використання скорингових моделей, при побудові яких виконується збір та попередня обробка даних, визначення ваг критеріїв за допомогою експертних та статистичних підходів. Скорингова карта будується із розрахунком загального скору характеристик даних для досліджуваних об'єктів [110].

Новизна цього підходу полягає в інтеграції різних моделей та методів в рамках єдиної інформаційної технології, в тому числі, методів штучного інтелекту та машинного навчання. За такого підходу критичні ситуації, ризики настання негативних подій можна відстажувати, аналізуючи великі обсяги даних

у режимі реального часу, виявляючи приховані закономірності та ранні ознаки потенційних проблем, що складно або й навіть неможливо дослідити використовуючи класичні методи аналізу. Крім того, можливість швидкого оновлення скорингових карт з урахуванням змін зовнішнього середовища забезпечує більш оперативне реагування на загрози та ризики критичних ситуацій.

Інтеграція даних з різних джерел, таких як системи моніторингу інформаційно-комунікаційної мереж, сенсори та історичні записи, дозволяє створити повну картину поточного стану об'єктів. Це відкриває можливості для оптимізації процесів управління та підвищення ефективності операційної діяльності.

Основна задача полягає у розробленні системи скорингових карт для оцінки ризиків об'єктів критичної інфраструктури з урахуванням даних про кібератаки, що накопичуються у системі кіберзахисту. Виконання даної задачі відбувається за наступних етапів:

1. Збір та попередня обробка даних (перевірка на повноти, імпуція пропущених значень, очищення та нормалізація).
2. Визначення ваг критеріїв (використання експертних оцінок (за методом аналізу ієрархій) та статистичних методів (регресійного аналізу) для об'єктивного визначення внеску кожного показника).
3. Побудова скорингової карти (розрахунок загального скору об'єкта як зваженої суми нормалізованих показників, що дозволяє ранжувати об'єкти за рівнем ризику).
4. Формулювання рекомендацій (розроблення системи управлінських впливів на основі отриманих скорингових значень для своєчасного та адекватного реагування на кіберзагрози).

Попередньо, задачу можна сформулювати наступним чином. Нехай, є набір даних, представлений матрицею X , розмірності $N \times M$ де N — кількість об'єктів, а M — кількість критеріїв оцінки (наприклад, показники тривалості атаки, масштабність, тип атаки тощо).

Для кожного показника j , що досліджується застосовують метод мін-макс нормалізації (2.19):

$$x_{ij}^{(n)} = \frac{x_{ij} - \min(x_i)}{(x_j) - \min(x_j)}, i = 1, \dots, N; j = 1, \dots, m \quad (2.19)$$

де x_{ij} — початкове значення, $\min(x_j)$ та (x_j) — мінімальне та максимальне значення для критерію j відповідно.

Ваги критеріїв визначають наступним чином. Нехай w_j — вага критерію j , яка визначається за допомогою методів аналізу ієрархій (АНР) або регресійного аналізу, з умовою (2.20):

$$\sum_{j=1}^m w_j = 1 \text{ та } w_j \geq 0 \quad (2.21)$$

Загальний скор для об'єкта i розраховується як зважена сума нормалізованих показників:

$$S_i \sum_{j=1}^m w_j x_{ij}^{(n)}, i = 1, \dots, N \quad (2.22)$$

Отримані значення S_i дозволяють ранжувати об'єкти за ступенем загрози.

На основі порогових значень скору, які поділяють об'єкти на категорії (наприклад, високий, середній та низький ризик), формується рекомендаційна функція (2.23):

$$R(S_i) = \{R_{\text{високий}}, S_i < T_1 \text{ } R_{\text{середній}}, T_1 \leq S_i < T_2 \text{ } R_{\text{низький}}, S_i \geq T_2 \} \quad (2.23)$$

де T_1 та T_2 — порогові значення скору, а $R_{\text{високий}}$, $R_{\text{середній}}$, $R_{\text{низький}}$ — відповідні рекомендації щодо заходів з управління ризиками.

Запропонована математична модель забезпечує інтеграцію різних типів даних та критеріїв оцінки в єдину систему, що дозволяє не лише ранжувати

об'єкти критичної інфраструктури за рівнем ризику, а й формувати конкретні управлінські рекомендації для своєчасного реагування на кіберзагрози.

Висновки до розділу 2

1. Визначені концептуальні та теоретичні засади підтримки прийняття рішень в управлінні регіональним розвитком як на рівні регіону, так і на рівні громади.
2. На підставі аналізу існуючих підходів розробки системи підтримки прийняття рішень, що використовуються різних сферах, для вирішення різноманітних задач запропонованих вітчизняними та закордонними вченими та фахівцями-практиками встановлено, що то існує значна кількість підходів до проектування та реалізації таких систем.
3. Встановлено, що враховуючи рівень сучасних інформаційних технологій, активно розвивається напрямок комп'ютерних систем підтримки прийняття рішень.
4. На основі системного аналізу проблеми сформульовані вимоги до інформаційно-аналітичного забезпечення та методології розв'язання задач підтримки прийняття рішень в управління регіональним розвитком.
5. Виходячи з особливості об'єкта дослідження на основі проведеного аналізу інформаційних джерел, запропоновано метод підтримки прийняття рішень на основі мережецентричного підходу, який на відміну від чинних, забезпечує взаємодію органів державного управління та територіальних громад у розподіленому інформаційному середовищі з підтримкою автоматизованого обміну даними та адаптації управлінських стратегій до зовнішніх і внутрішніх умов, що збільшить гнучкість і швидкість реагування на зміни ситуації
6. Для вирішення задачі забезпечення ефективної прийняття рішень щодо забезпечення взаємодії складових системи управління регіональним

розвитком на різних рівнях управління, запропоновано формалізовану методологію підтримки прийняття рішень в управлінні регіональним розвитком, яка, на відміну від відомих охоплює повний цикл від збору та інтеграції різнорідних структурованих і неструктурованих даних до формування оптимального рішення, та базується на системному підході з використанням багатомодельного і багатокритеріального аналізу, інтероперабельної мережецентричної архітектури й інтелектуальних технологій опрацювання даних, що дозволяє забезпечити узгодженість стратегій розвитку та підвищити обґрунтованість управлінських рішень.

7. Інформаційну взаємодію учасників процедур підтримки прийняття рішень і управлінні регіональним розвитком, запропоновано реалізувати в рамках єдиної інформаційної технології, в основу якої покладено інтеграцію принципів системного аналізу, методів обробки та оцінювання якості даних і прогнозного моделювання з використанням нових математичних моделей та їх композицій, що збільшить достовірність результатів і адаптивність до різних предметних областей.
8. Розвинуто метод оцінювання ризиків і невизначеностей на основі процедури оцінювання варіантів розвитку ситуації, який, на відміну від існуючих, дозволяє враховувати невизначеності та ризики різних типів, що забезпечує підвищення точності оцінок і надійності остаточних управлінських рішень.
9. У роботі пропонується системне використання уніфікованого набору інформаційних технологій, об'єднаних у єдиній інформаційній системі регіонального розвитку, побудованій за мережецентричним принципом на основі інтероперабельності, застосування комплексної системи захисту інформації з підтвердженою відповідністю

РОЗДІЛ 3

ІНФОРМАЦІЙНО-АНАЛІТИЧНЕ ЗАБЕЗПЕЧЕННЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ

3.1 Особливості аналітичного забезпечення регіонального рівня

Під час розв'язання задач прийняття рішень в управлінні розвитком соціально-економічних систем проблемним залишається питанням збору та оброблення даних, необхідних для аналізу стану та динаміки досліджуваних систем для подальшого їх використання у відповідних системах підтримки прийняття рішень. Обробка даних, що описують соціально-економічні системи (але, важливо враховувати й екологічну складову) як правило, супроводжується невизначеностями різного типу. Зокрема, це невизначеності структурні, статистичні й параметричні, але, враховуючи умови військового конфлікту, все частіше розглядається системна невизначеність

Основою інформаційного забезпечення дослідження регіонального розвитку є статистична інформація, що збирається Державною службою статистики України. Ці набори даних є структурованими, їх збір, оброблення та оприлюднення визначаються Законом України "Про офіційну статистику" [1]. Реформування регіонального рівня державного управління, вносить корективи і до порядку формування статистичної інформації. Так, з 01.01.2025 було прийнято нову Статистичну класифікацію територіальних одиниць України (NUTS-4) [111], яка змінює статистичну класифікацію територій України відповідно до норм європейського законодавства, і відповідно, й до методики збору та подання статистичної інформації. Що, звичайно ускладнює ретроспективний аналіз ситуації. Іншими джерелами даних є дані Національного банку України, міністерств, обласних, районних адміністрацій, тощо, які оприлюднені на офіційних сайтах. Крім внутрішніх інформаційних джерел для дослідження регіонального розвитку використовуються й дані міжнародних організацій, оприлюдненій у встановленому порядку. Крім того, громадські та

неурядові організації оприлюднюють власні спостереження переважно у вигляді текстових звітів. Звичайно, значну частину даних можна отримати з різних Інтернет-джерел. Тому питання збору та попередньої обробки даних, що характеризують процеси регіонального розвитку є актуальною задачею.

Для аналізу стану та динаміки національної соціально-економічної системи країни та регіонів відбувається порівняння прогностичних показників та їх виконання, планових показників та їх виконання. Щорічні прогностичні макропоказники розробляються та схвалюються постановою Кабінету Міністрів. Вони є основним орієнтиром щодо реалізації державної політики соціально-економічного розвитку. Інші нормативні документи визначають порядок реалізації державної регіональної політики. Це постанови Кабінету Міністрів України від 4 серпня 2023 р. № 816 «Деякі питання розроблення регіональних стратегій розвитку і планів заходів з їх реалізації та проведення моніторингу реалізації зазначених стратегій і планів заходів» [120], і від 4 серпня 2023 р. № 817 «Деякі питання розроблення Державної стратегії регіонального розвитку України і плану заходів з її реалізації та проведення моніторингу реалізації зазначених Стратегії і плану заходів».

Постановою Кабінету Міністрів України від 15.03.2024 №305 «Деякі питання проведення моніторингу та оцінювання державної регіональної політики» врегульовано основні засади оцінювання результативності державної регіональної політики.

Для оцінювання складових соціально-економічного розвитку держави використовується «Концепція створення системи рейтингової оцінки регіонів, галузей національної економіки, суб'єктів господарювання» []. схвалена розпорядженням Кабінету Міністрів України від 1 квітня 2004 р. № 208-р [111]

Однак за сучасних динамічних стрибкоподібних зміни економічної ситуації, що характеризується наявністю невизначеностей та ризиків, значним впливом людського фактору, дослідження значно ускладнюється і методи економіко-математичного та статистичного аналізу не дають прийнятних результатів. Основним і суттєвим недоліком таких підходів є обмеженість часового інтервалу

дослідження (1-3 роки). Що значно ускладнює дослідження циклічності та тенденцій зміни показників, ускладнює урахування багатofакторних залежностей за тривалий період, розкриття невизначеностей.

Сучасні тенденції економічного розвитку ставлять задачі, які потребують аналізу великих обсягів даних, пошуку закономірностей та зв'язків, часто неявних, що ускладнює їх оцінювання виключно методами економетрики та статистики, як це закладено у більшості державних методик. Наприклад, значна кількість статистичних даних у зв'язку із змінами у статистичній методології або періодично не збирається, або збиралась не постійно, що не дозволяє сформувати часові ряди показників навіть за 5-10 років. Така ситуація є причиною неможливості чіткого та повного визначення кількісних характеристик більшості економічних процесів, дослідження регіонального розвитку, міжгалузевого обміну, забезпечення узгодженості дій виконавців та координаторів процесів на різних рівнях державного та регіонального управління. Процес прийняття рішення ускладнюється, виникає потреба застосування інструментів системного аналізу, розроблення нових аналітичних підходів та інформаційних технологій [115].

Вирішенню проблеми забезпечення повноти вхідної інформації на етапі збору та попередньої обробки присвячено багато робіт вітчизняних та закордонних фахівців [6-17], однак, вона залишається невирішеною. Відсутня також й ефективна інформаційна технологія аналізу та обробки такої інформації.

3.2 Моделі та методи попередньої обробки даних

Для подолання можливих невизначеностей статистичного, структурного та параметричного типів запропоновано множину методів, залежно від типу невизначеності. Запропонована методика попереднього опрацювання даних вирізняється тим, що при формуванні масивів вхідних даних їх інформативність, синхронність та коректність забезпечуються на етапі попередньої обробки, яка виконується для приведення їх до форми, що забезпечить можливість коректного

застосування методів оцінювання параметрів моделі та отримання їх статистично значущих оцінок. Зокрема, запропоновано використовувати методику корегування значних імпульсних (екстремальних) значень, нормування вимірів у заданих межах, логарифмування великих значень та фільтрації шумових складових, виявлення та усунення мультиколінеарності [52, 56-61].

Особливу увагу у даній роботі приділено дослідженню вхідного набору даних, оскільки необхідно забезпечити як репрезентативність, так і варіабельність наборів даних, враховуючи при цьому, що значна кількість факторів повинна бути врахована виходячи з вимог нормативних документів, що може бути причиною мультиколінеарності вхідних змінних.

Для підготовки даних можна застосувати кілька методів, які забезпечать їхню чистоту, однорідність і придатність для подальшого аналізу. Серед них варто виділити методи перевірки на повноту даних, які включають аналіз наявності пропущених значень та їх корекцію шляхом використання стратегій заповнення – середнього, медіани або методів машинного навчання, наприклад, KNN-імпутації. Очищення даних також є критичним етапом: видалення або корекція аномалій і дубльованих записів допомагає уникнути спотворення результатів аналізу.

Нехай наш набір даних представлений матрицею

$$X = \{x_{ij}\}, i = 1, \dots, N; j = 1, \dots, p \quad (3.1)$$

де N – кількість спостережень, а p – кількість характеристик. Для кожного елемента x_{ij} можна визначити індикатор відсутності даних:

$$I(x_{ij}) = \{1, \text{якщо } x_{ij} \text{ відсутнє } 0, \text{якщо } x_{ij} \text{ присутнє} \quad (3.2)$$

Імпутація за середнім значенням передбачає наступні дії. Для кожного стовпця j , де x_{ij} не пропущені, обчислюють середнє значення:

$$\underline{x}_j = \frac{1}{N_j} \sum_{i: I(x_{ij})=0} x_{ij} \quad (3.3)$$

де N_j – кількість спостережень без пропущених значень для ознаки j . Пропущене значення x_{ij} замінюється на $\underline{x}_j$.

Для імпутації за медіаною, медіану для ознаки j позначили як:

$$\underline{x}_j = \text{median}(\{x_{ij}: I(x_{ij}) = 0\}) \quad (3.4)$$

І, відповідно, замінити пропущені значення на $\underline{x}_j$.

За KNN-імпутації, приймаємо, що нехай для спостереження i з пропущеним значенням в ознаці j ми обираємо k найближчих спостережень, для яких x_{lj} не пропущене. Визначимо відстань $d(i, j)$ між спостереженнями (наприклад, за Евклідовою відстанню) з використанням інших ознак, де дані не відсутні. Тоді імпутоване значення можна обчислити як зважене середнє:

$$x_{ij}^* = \frac{\sum_{l \in N_k(i)} w_{il} * x_{li}}{\sum_{l \in N_k(i)} w_{il}} \quad (3.5)$$

де $N_k(i)$ – множина k найближчих сусідів, а ваги w_{il} можуть бути обернено пропорційні відстані, наприклад, $w_{il} = \frac{1}{d(i, l)}$.

Видалення дубльованих записів здійснюється наступним чином. Нехай два спостереження x_i та x_j вважаються дубльованими, якщо $x_i = x_j$

При виявленні таких записів один із них видаляється.

Для виявлення аномалій за допомогою Z-score, для кожної ознаки j обчислюють середнє μ_j та стандартне відхилення σ_j . Потім для кожного значення розраховують

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (3.6)$$

Якщо $|z_{ij}| > \theta$, де θ – заданий поріг, наприклад, 3, то значення x_{ij} можна вважати аномальним. Аномальні значення можуть бути або вилучені, або скориговані за допомогою відповідних стратегій.

Ці математичні постановки забезпечують формалізацію підходу до підготовки даних, дозволяючи оцінити їх повноту, провести імпутацію пропущених значень, а також очистити дані від дубльованих записів та аномалій, що в сукупності створює надійну основу для подальшого аналізу та побудови скорингових карт для оцінки ризиків.

Нормалізація значень забезпечує можливість коректного порівняння різних показників, що мають різні діапазони або одиниці вимірювання. Для цього можна використовувати масштабування за методом мін-макс або стандартизацію за Z-score, що дозволить привести дані до спільної шкали. При роботі з часовими мітками важливо привести їх до єдиного формату, а також за необхідності виділити додаткові ознаки, як-от день тижня чи година, які можуть впливати на характеристику атаки.

Також корисними є методи обробки категоріальних даних, наприклад, one-hot encoding або label encoding, що дозволяють перевести текстові або номінативні змінні у числовий формат для подальшого аналізу. Якщо дані мають розподіл з високою асиметрією або містять викиди, можна застосувати логарифмічні перетворення або Box-Cox трансформації, які допоможуть зменшити вплив екстремальних значень. Всі ці методи разом забезпечують надійну основу для подальшої побудови скорингової карти та аналізу ризиків.

При визначенні ваг для кожного критерію, що використовуються в скоринговій системі, можна скористатися як експертними оцінками, так і статистичними методами, що забезпечують об'єктивність визначення внеску кожного показника у загальний результат. Одним із поширених підходів є метод аналізу ієрархій (АНР). За цим методом експерти проводять попарне порівняння критеріїв, що дозволяє сформувати матрицю порівнянь $A = [a_{ij}]$, де кожен елемент a_{ij} відображає відносну важливість критерію i порівняно з критерієм j . Властивості цієї матриці включають такі умови: $a_{ij} = 1$ для всіх i та симетрію

за принципом оберненості, тобто $a_{ji} = 1/a_{ij}$. Далі знаходиться власний вектор w , що задовольняє рівняння (3.7):

$$Aw = \lambda_{max} w \quad (3.7)$$

де λ_{max} – максимальне власне значення матриці Aw . Після нормалізації цього вектора, тобто з умовою (3.8):

$$\sum_{i=1}^m w_i = 1 \quad (3.8)$$

отримуємо ваги $w_1, w_2, \dots, w_m$, які відображають відносну важливість кожного критерію.

Іншим підходом є статистичний метод, коли наявні історичні дані про результати оцінки або ефективність роботи об'єктів дозволяють застосувати регресійний аналіз. Припустимо, що для кожного об'єкта ми маємо значення $x_1, x_2, \dots, x_m$ за відповідними критеріями та спостерігаємо кінцевий результат Y . Тоді можна побудувати модель:

$$Y = \beta_1 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \epsilon \quad (3.9)$$

де коефіцієнти $\beta_1, \beta_2, \dots, \beta_m$ вказують на внесок кожного критерію у визначення результату. Нормалізуючи ці коефіцієнти (наприклад, поділивши кожен на суму абсолютних значень усіх коефіцієнтів), можна отримати ваги, які відображають важливість критеріїв з точки зору їх впливу на кінцевий показник.

Пропонована методика попереднього аналізу даних, призначена для визначення наявності зв'язків та аномалій для їх урахування під час побудови моделей процесів, що характерні для об'єктів соціально-еколого-економічних систем складається з таких етапів.

Етап 1. Графічний аналіз даних із використанням діаграм розсіювання включає використання графіків та кореляційного аналізу.

Етап 2. Діагностика екстремальних значень та викидів, включає аналіз

діаграм розсіювання та графіків на наявність викидів, за потреби, з'ясовуються причини появи викидів.

Етап 3. Аналіз впливу екстремальних значень, студентизовані залишки; D-статистика Кука; статистика DFFITS; статистика DFBETAS.

Етап 4 Дослідження даних на наявність колінеарності, включає використання VIF-статистики, індекс стану та пропорцію варіації. Перевірка пропорції варіації включає дослідження студентизованих залишків, D-статистика Кука та DFFITS статистика.

Етап 5. Аналіз категоріальних даних, що включає дослідження їх розподілу, проведення попереднього дослідження наявності зв'язку між змінними (використовується статистика χ^2 -квадрат Пірсона, V-статистика Крамера, співвідношення шансів, χ^2 -квадрат тест Мантел-Гаензеля)

Етап 6. Стандартизація даних використовує наступні методи стандартизації. В загальному випадку для стандартизації даних, в тому числі і при заповненні пропусків, саме стандартизованими значеннями, використовується наступна формула [26, 68] (3.10):

$$result = add + multiply \times \frac{(original - location)}{scale} \quad (3.10)$$

де *result* – кінцеве значення (стандартизоване), *add* – адитивна константа (зміщення), *multiply* – мультиплікативна константа, *original* – початкове (оригінальне) значення (до стандартизації), *location* – оцінка вибіркового середнього значення (як правило математичне сподівання), *scale* – оцінка вибіркового значення масштабу (як правило стандартне відхилення).

Середнє мінімального інтервалу - обчислюється за спеціальним алгоритмом [21, 69]. Для обчислення вибірових статистик використовується 90% вибірки, 10% вибірки з лівого і правого боку розподілу виключається з розрахунків. За своєю суттю це вибіркова статистика за усіченою вибіркою, що очищена від викидів.

Відстань Евкліда між двома точками, в загальному випадку обчислюється за формулою:

$$\rho(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} ,$$

але для випадку однієї змінної вироджується в формулу:

$$\rho(x) = \sqrt{\sum_{i=1}^n (x_i)^2} .$$

В загальному випадку, відстань Мінковського [70] між двома точками обчислюється за формулою:

$$\rho(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}$$

але для випадку однієї змінної вироджується в:

$$\rho(x) = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

важливо перевірити наявність викидів в даних.

Спостереження, що знаходяться відчутно далеко від основної частини даних, є викидами. Ці спостереження часто є помилками даних або відображають незвичайні обставини, що виникли в процесі функціонування об'єкта. У будь-якому випадку, коректним є підхід, коли окремо з'ясовують причини виникнення викидів.

Для діагностики екстремальних значень використовуються такі статистики [21, 52, 59, 60]: студентизовані залишки; RSTUDENT-залишки; статистика Cook's D; статистика DFFITS; статистика DFBETAS.

Один із підходів, щодо перевірки викидів – використання студентизованих залишків. Вони обчислюються шляхом ділення залишків на значення стандартної похибки. Для моделі, яка добре підходить до даних і не має відхилень, більша частина студентизованих залишків повинна бути близька до 0.

В цілому, студентизовані залишки, що мають абсолютне значення менше 2,0, інколи виникають випадково. Студентизовані залишки, модуль яких знаходиться між 2,0 та 3,0, спостерігаються не часто і дійсно, можуть бути викидами. Модуль значення студентизованих залишків перевищує значення 3,0, зустрічаються ще рідше і підлягають більш глибокому вивченню та дослідженню.

Якщо студентизовані залишки $|SR| > 2$, то є підозра на наявність викиду, якщо $|SR| > 3$ – однозначно, є викид, необхідно більш глибоке вивчення та дослідження даних.

Студентизовані залишки є звичайними залишками, розділеними на їх стандартні похибки. RSTUDENT залишки схожі на студентизовані залишки, за винятком того, що вони обчислюються після видалення i -го екстремального спостереження. Інший варіант пояснення, RSTUDENT залишки – це різниця між реальним значенням відгуку Y та значенням-прогнозу відгуку $\hat{Y}$, за виключенням екстремального значення.

D- статистика Кука використовується для ідентифікації екстремальних значень. Ця статистика вимірює зміну оцінок параметрів, що виникає при видаленні кожного спостереження (3.11) [9, 16]:

$$D - \text{статистика Кука} = \left(\frac{1}{ps^2} \right) (b - b_{(i)})' (X'X) (b - b_{(i)}), \quad (3.11)$$

де p – кількість параметрів регресії; s^2 – середньоквадратична похибка регресійної моделі; b – вектор оцінок параметрів; $b_{(i)}$ – вектор оцінок параметрів, отриманих після видалення i -го спостереження; $X'X$ – скоригована сума квадратів і матриці перехресних добутків.

В якості порогу відсікання, для ідентифікації екстремальних значень пропонується використовувати (3.12) [9, 16]:

$$D - \text{статистика Кука}_i > \frac{4}{n}, \quad (3.12)$$

де n – кількість спостережень.

Статистика DFFITS (3.13) вимірює вплив, який має i -те спостереження на значення, що прогнозується [9, 16, 21].

$$DFFITS_i = \frac{\hat{Y}_i - \hat{Y}_{(i)}}{s(\hat{Y}_i)} \quad (3.13)$$

де $\hat{Y}_i$ – i -те значення прогнозу, $\hat{Y}_{(i)}$ – i -те значення прогнозу, при видаленні i -го спостереження, $s(\hat{Y}_i)$ – стандартна похибка i -го значення прогнозу.

Великі значення статистики DFBETAS вказують на те які саме регресори можуть бути причиною екстремального впливу.

За [26], прийнято значення порогу рівного 2, при перевищенні якого вважається, що спостереження екстремальне, або обчислювати поріг як $2\sqrt{\frac{1}{n}}$, тобто з поправкою на розмір вибірки.

Виявлення викидів за графіками можливо на невеликих вибірках, але для великих наборів даних має сенс аналізувати детально саме тільки підмножини екстремальних значень.

Метою побудови будь-якої математичної моделі є пошук найкращої площини, гіперплощини або огибаючих кривих, що проходять через дані для прогнозування відгуку.

Проблема колінеарності полягає в тому, що:

1. Колінеарність може збільшити уявну статистичну значущість ефектів, за рахунок повторного включення тієї ж порції інформації.
2. Колінеарність прагне збільшити дисперсію оцінок параметрів i , отже, збільшити похибку прогнозування.

За наявності колінеарності оцінки коефіцієнтів моделі нестійкі, вони мають велику дисперсію. Більше того, реальний зв'язок між Y та X_S може сильно відрізнятись від отриманих результатів. При цьому наявність колінеарності не порушує умов, що накладається на лінійну регресію.

Для діагностики колінеарності використовується набір інструментів для кількісного оцінювання проблеми колінеарності та визначення множини

регресорів X_S , що містять колінеарність.

Метрика VIF, скорочення від англійської Variance Inflation Factor, являє собою міру величини колінеарності. VIF – це відносна міра росту дисперсії внаслідок наявності колінеарності (3.14) [16, 21, 24, 26].

$$VIF_i = \frac{1}{1 - R_i^2} \quad (3.14)$$

де R_i^2 – статистика R-квадрат для регресорів X_i на всіх регресорах моделі X_S .

Наприклад, якщо модель для прогнозування відгуку Y включає регресори X_1, X_2, X_3, X_4 (тобто X_S для $i = 1, \dots, 4$), то для обчислення статистики R квадрат на регресорі X_3 , будується модель, що включає регресори X_1, X_2, X_4 . Обчислене значення статистики R^2 підставляють до формули 3.6:

$$VIF_3 = \frac{1}{1 - R_3^2},$$

якщо значення $VIF_i > 10$, то присутня колінеарність.

Значення VIF_i обчислюється для кожного регресора моделі.

Колінеарність може істотно вплинути на результат побудови регресії, що будується за методом STEPWISE (поеапного включення регресорів)[26-35]. Оскільки значимість важливих змінних може бути замаскована колінеарністю, як наслідок кінцева модель може не включати дуже важливі регресори. Тому перевірку на колінеарність треба виконувати перед побудовою моделі в автоматичному режимі.

В загальному випадку існують інші підходи щодо вирішення проблеми колінеарності, наприклад, два методи – гребнева регресія та регресія на головних компонентах [44-50, 65]. Крім того, центрування змінних-предикторів іноді може дозволити вирішити проблему колінеарності, особливо для випадків поліноміальної регресії в ANCOVA (Analysis of covariance) моделях.

Пропонована методика попередньої обробки даних включає виконання наступних етапів.

Етап 1. Попередній аналіз: використання описових статистик, графіків та кореляційного аналізу.

Етап 2. Вибір моделі-кандидата: ідентифікація однієї або декількох моделей-кандидатів.

Етап 3. Валідація: побудова та аналіз графіків та діаграм розсіювання залишків у порівнянні з значеннями прогнозів, перевірка рівності дисперсії.

Етап 4. Виявлення колінеарності та екстремальних спостережень: перевірка на колінеарність з використанням VIF статистики, індексу стану та пропорції варіації (обчислення студентизованих залишків, D-статистика Кука та DFFITS статистик).

Етап 5. Тестування моделі: якщо на етапах 3 та 4 очевидно є необхідність перевірки моделі, то будується нова модель із поверненням до цих кроків.

Етап 6. Тестування прогнозів: якщо є можливість, то модель перевіряється на даних, що не використовувалися для її побудови.

Коли дані неповні, спотворені або є сумніви щодо їх достовірності, для отримання прогнозів розвитку процесів різної природи, одним із шляхів вирішення задачі інформаційної невизначеності є застосування методики, основаної на виявленні аналогій на основі аналізу процесів за подібністю. Суть підходу полягає у пошуку часових рядів зі схожими статистичними поведінками, що допомагає зробити висновки про наявність поведінкових шаблонів.

Представлена методика дозволяє порівняти два часові ряди з різною довжиною використовуючи метод динамічного перетворення, а також усереднення (згладжування) значень. Для врахування різної поведінки рядів у часі, під час аналізу подібності вимірюється відстань між першою часовою послідовністю та другою із урахуванням впорядкування. Зсув часу (ковзання) послідовно зміщує перший часовий ряд уздовж часової шкали, з метою пошуку подібних зразків, які можуть виникати в довільні моменти часу. Перетворення часового ряду дозволяє розтягувати або стискати часовий вимір, для випадків незначної відмінності у часі між схожими шаблонами. Спеціалізовані методи та підходи щодо перетворення часового ряду за часовою шкалою дозволяють розширювати та стискати цикли з метою виявлення схожих форм сезонного циклу, навіть при наявності незначних відмінностей. Загальний опис методу

динамічного перетворення часового ряду в часі можна представлений в роботі [61].

Реалізація пропонованої методики виконується після попередньої обробки даних із виконанням таких процедур як: агрегування даних (перетворення транзакційних даних до необхідних часових вимірів), обробка пропусків даних, завдання сезонного циклу (погодинного, добового, тижневого, місячного, p-річного, тощо).

Встановлено, що для підготовки даних для побудови математичних моделей у задачах моделювання розвитку соціально-економічних систем доцільно використовувати підхід, що ґрунтується на розділенні всієї вибірки вхідних даних на тренінгову, валідаційну та тестову. Проблемним питанням побудови математичних моделей нелінійних нестаціонарних процесів є так зване «прокляття розмірності», коли необхідно будувати модель на коротких часових вибірках даних. Тобто використання існуючих методів прогнозування в задачах прийняття рішень, які ґрунтуються на аналітичних процедурах, логічних правилах та раціональному експертному оцінюванні, у багатьох випадках не дає бажаного результату стосовно якості оцінок прогнозів, а тому в процесі дослідження зазвичай вирішується проблема системного використання альтернативних методів моделювання і прогнозування для підвищення якості оцінок прогнозів для кожного конкретного випадку. Зокрема, для моделей процесів, які описують функціонування соціально-економічних систем, важливо знайти підмножину вхідних змінних (які називаються ознаками або атрибутами), що якнайкраще характеризуватимуть предметну область. Для цього застосовуються три стратегії – стратегія фільтра (наприклад, накопичення ознак), стратегія обгортання (наприклад, пошук згідно заданої точності) і стратегія вкладення (вибираються ознаки для додавання або видалення в міру побудови моделі, заснованої на аналізі похибок прогнозування).

Для застосування техніки відбору ознак при побудові моделей процесів, що мають місце у системі захисту об'єктів соціально-економічних систем використано такі підходи: спрощення моделей, щоб зробити їх простішими для

інтерпретації дослідниками / користувачами; зменшення часу на тренування (навчання) моделі; зменшення розмірності вхідних даних; покращене узагальнення шляхом уникнення перенавчання (формально, це зменшення дисперсії).

Коли дані неповні, спотворені або є сумніви щодо їх достовірності, для отримання прогнозів розвитку процесів різної природи, одним із шляхів вирішення задачі інформаційної невизначеності є застосування методики, основаної на виявленні аналогій на основі аналізу процесів за подібністю. Суть підходу полягає у пошуку часових рядів зі схожими статистичними поведінками, що допомагає зробити висновки про наявність поведінкових шаблонів.

Представлена методика дозволяє порівняти два часові ряди з різною довжиною використовуючи метод динамічного перетворення, а також усереднення (згладжування) значень. Загальний опис методу динамічного перетворення часового ряду в часі можна представлений в роботі [61].

Реалізація пропонованої методики виконується після попередньої обробки даних із виконанням таких процедур як: агрегування даних (перетворення транзакційних даних до необхідних часових вимірів), обробка пропусків даних, завдання сезонного циклу (погодинного, добового, тижневого, місячного, p-річного, тощо).

Для подолання можливих невизначеностей статистичного, структурного та параметричного типів запропоновано множину методів, залежно від типу невизначеності. Запропонована методика попереднього опрацювання даних вирізняється тим, що при формуванні масивів вхідних даних їх інформативність, синхронність та коректність забезпечуються на етапі попередньої обробки, яка виконується для приведення їх до форми, що забезпечить можливість коректного застосування методів оцінювання параметрів моделі та отримання їх статистично значущих оцінок. Зокрема, розроблено методику корегування значних імпульсних (екстремальних) значень, нормування вимірів у заданих межах, логарифмування великих значень та фільтрації шумових складових, виявлення та усунення мультиколінеарності [52, 56-61].

Для застосування техніки відбору ознак при побудові моделей використано такі підходи, як спрощення моделей, зменшення розмірності вхідних даних; покращене узагальнення шляхом уникнення перенавчання (формально, це зменшення дисперсії).

3.3 Методика застосування моделі різних типів для аналізу та моделювання процесів регіонального розвитку

Дослідження економіки регіонів України, в тому числі й в період різних кризових явищ, економічного зростання, проведення соціально-економічних та територіальних реформ засвідчили чутливість економічної системи країни до змін, що відбуваються у світовій економіці за достатньо високої стійкості регіонів, особливо аграрних, у протидії кризовим явищам. Коректне планування майбутнього розвитку соціально-економічних та фінансових процесів ґрунтується на оцінках прогнозів, обчислених за допомогою відповідних (адекватних) математичних і статистичних моделей [1, 2]. Для розв'язання задач коротко- і середньострокового прогнозування сьогодні застосовують методи, що ґрунтуються на різних ідеях стосовно побудови прогнозуючих моделей [2, 3]. Найбільш популярними є відносно прості у побудові та подальшому використанні моделі множинної регресії, авторегресії з ковзним середнім (АРКС), нейронні мережі, нейронечіткі структури, моделі на основі нечіткої логіки, регресія на опорних векторах, ймовірнісні моделі байєсівського типу та інші [3]. У випадку наявності у модельованих процесах невизначеностей статистичного і структурного характеру прийнятні результати стосовно оцінок прогнозів можна отримати за допомогою нечіткої логіки та ймовірнісних моделей, зокрема байєсівських мереж (БМ) [4]. БМ – ймовірнісна модель у вигляді спрямованого ациклічного графа, вершинами якого є вхідні та вихідні змінні процесу, а дуги вказують на наявні причинно-наслідкові зв'язки між вибраними змінними. Кожна змінна характеризується таблицею умовних ймовірностей, яка формується на основі апріорної інформації про ці змінні у

вигляді експертних оцінок і статистичних даних. Класифікація моделей представлена на рис.3.1.

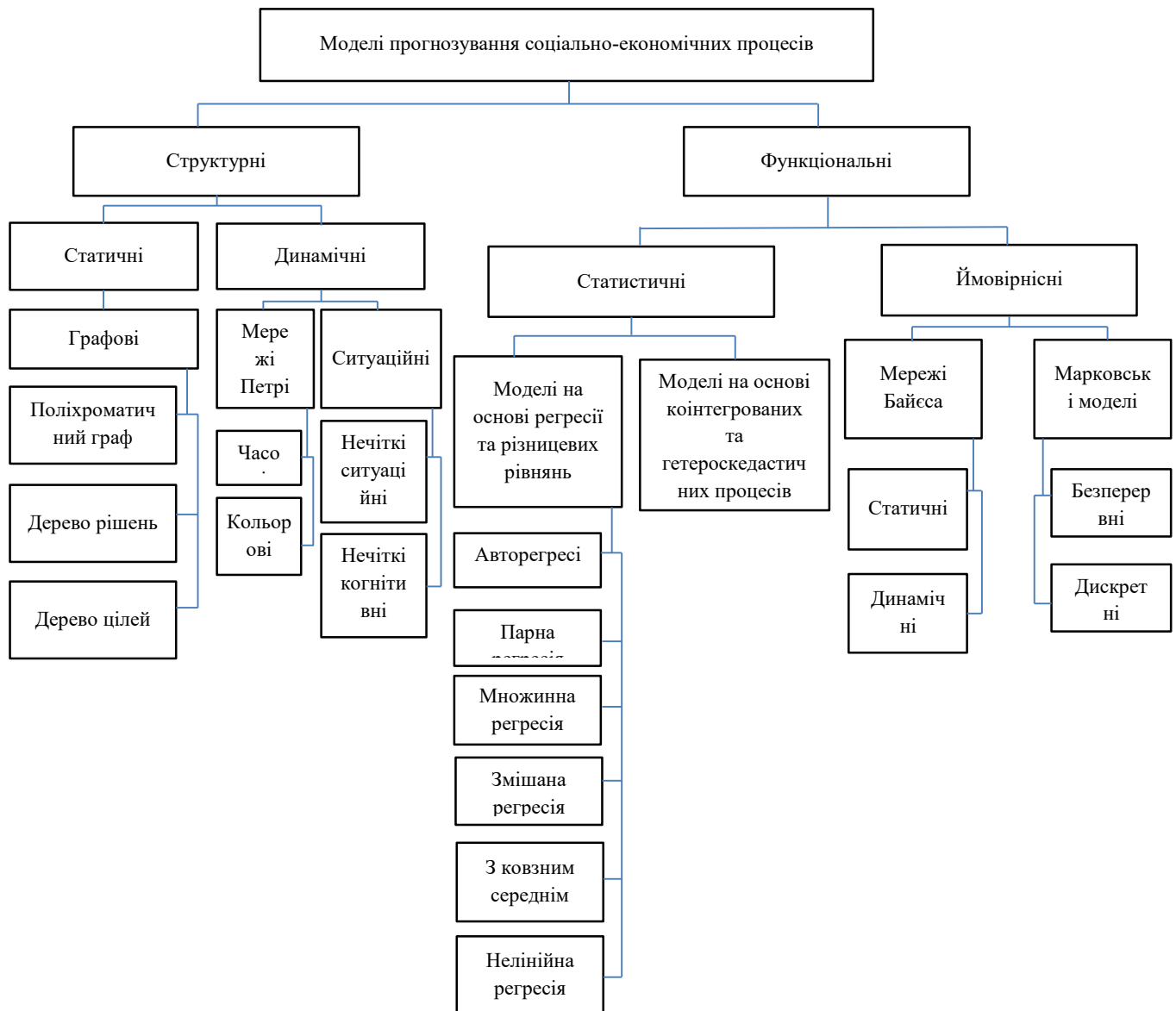


Рисунок 3.1 – Класифікація моделей соціально-економічних процесів [1]

Моделі та методи штучного інтелекту сьогодні стають все більш поширеною складовою компютерних систем підтримки прийняття рішень. Машинне навчання (Machine Learning) - напрям в інформатиці, який за традицією називають “штучний інтелект” [119-122]. Його метою є створення алгоритмів самонавчання на основі аналізу емпіричних даних. Машинне навчання стає все більш поширеним інструментом. Воно використовується: для класифікації повідомлень електронної пошти; отримання знань про систему

оцінювання користувача та підбору рекомендацій для нього, що ґрунтуються на цій інформації; визначення оптимального контенту для залучення потенційних клієнтів; встановлення ймовірності виграшу справи та відповідності юридичним нормам пред'явлених рахунків, все частіше його використовують і у системах кіберзахисту, системах захисту об'єктів критичної інфраструктури [122].

Інша група методів - навчання з учителем - це методи машинного навчання, які передбачають керування процесом навчання безпосередньо на основі заданого цільового значення. Моделі з учителем керуються похибкою між отриманим значенням та заданим, метою навчання є мінімізація похибки. Навчання без учителя дещо схоже з кластерним аналізом [119-125].

Ансамблі навчання - множина моделей, застосовувана для прогнозування. За рахунок побудови значної кількості моделей, отримують сукупність алгоритмів, що разом дають кращий результат, ніж кожна модель окремо.

В основу еволюційних та генетичних алгоритмів (Evolution Analysis, Genetic Algorithms) покладено такі принципи, як мутації, успадкування від батьківських елементів, природний відбір за певними ознаками. Методика алгоритмів представлена на основі успадкування від батьківських вузлів найкращих рис або мутації, для покращення якості всієї моделі. Як і за природного відбору, залишається найбільш сильний та «пристосований до життя» елемент.

Нейронні мережі – специфічний вид моделей, що складається з багатьох шарів навчання та призначений для вирішення широкого спектру задач, які оперують з великими обсягами даних.

Для унаочнення залежностей в даних, використовують графічні методи подання результатів аналізу. Зображення результатів у вигляді схем, таблиць, зображень, діаграм сприяє розумінню взаємозв'язків, особливостей даних та допомагає краще інтерпретувати результати дослідження [119, 126]. Новими методами візуалізації даних є такі, як хмара тегів, кластерограма, просторовий потік.

Складні задачі прогнозування часових рядів досить часто вирішуються за допомогою використання методів штучних нейронних мереж (ANN) [119, 126].

Особливістю застосовування ANN є те, що вони здатні до нелінійного моделювання, де кожна модель є адаптивною та формується в залежності від поточних даних.

Найпростіша ANN-модель, як правило, має три шари (англ. layers): вхідний, прихований та вихідний. Якщо говорити про складніші штучні нейронні мережі, то вони матимуть більше одного прихованого шару, які з'єднані ациклічними зв'язками [119, 126]. Математична формалізація вихідного значення нейронної мережі має наступний вигляд [119, 126]:

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g(\beta_{0j} + \sum_{i=1}^p \beta_{ij} y_{t-i}) + \varepsilon_t, \forall t \quad (3.15)$$

де p – кількість вхідних змінних;

q – кількість прихованих шарів;

α_j, β_{ij} - вагові коефіцієнти;

ε_t – випадковий шум.

Як функцію g використовують наступні [8]:

1) сигмоїдна функція (також інколи зветься логістичною):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.16)$$

Дана функція є однією із найпоширеніших функцій активацій, яка використовуються у машинному навчанні.

2) гіперболічний тангенс:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.17)$$

Ця функція, у свою чергу, для багатошарових нейронних мереж дає більш кращий результат у порівнянні із сигмоїдною функцією. Функція гіперболічного

тангенсу є центрованою до нуля, що спрощує роботу у процесі зворотного поширення похибки [119, 126].

3) функція softmax:

$$f(x) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (3.18)$$

Найчастіше функція softmax використовується з метою визначення імовірності розподілу вектора дійсних чисел. Вона набуває значень у діапазоні від 0 до 1 [119, 126].

4) функція ReLU:

$$f(x) = \max(0, x) \quad (3.19)$$

Дана функція характеризується найбільшою точністю, саме тому є широко використовуваною. Функція ReLU наближається до лінійної функції, саме тому вона має властивості, які притаманні лінійним моделям. Це дозволяє досить легко використовувати методи градієнтного спуску для оптимізації моделей. Головною перевагою функції є те, що вона не потребує складних обчислень, як експонента чи ділення, що дозволяє зменшити час її виконання [119, 126].

Звичайно, що сьогодні існує багато функцій активації, які можуть бути використані у процесі побудови нейронної мережі, тому представлені найбільш часто використовувані.

Безпосередній процес навчання здійснюється за рахунок того, що функція похибки постійно мінімізується під час реалізації методу зворотного поширення помилки. Середньоквадратична функція похибки є цільовою функцією. Задачею є знаходження оптимальних значень ваг штучної нейронної мережі з метою мінімізації цільової функції [119, 126].

Найпоширенішими методами оптимізації є [119]:

1. Метод градієнтного спуску

$$x_{t+1} = x_t - \alpha f^t(x_t) \quad (3.20)$$

Відповідно до методу, з метою мінімізації функції необхідно виконати кроки, які є пропорційними протилежному значенню градієнта. Параметр α – це швидкість спуску. Чим більше значення α , тим більша ймовірність здійснення великих кроків з метою визначення найменшого значення функції. Проте у такому випадку існує ризик пропустити точку мінімуму. У іншому випадку – при низькому рівні швидкості навчання – алгоритм буде поступово йти у напрямку оптимального значення, проте реалізація, як правило, займає досить багато часу [119].

2. Метод стохастичного градієнтного спуску

Такий тип градієнтного спуску на кожній ітерації опрацьовує одне значення, яке використовується для навчання. Як результат, параметри моделі змінюються навіть після однієї ітерації. Це надає можливість знайти оптимальне значення цільової функції швидше, ніж при використанні методу звичайного градієнтного спуску. Проте, у випадку великої кількості даних, кількість ітерацій буде відповідно збільшуватися, що, як результат, вплине на час навчання нейронної мережі [119].

3. Метод градієнтного спуску міні-серіями

Цей метод працює швидше ніж звичайний або стохастичний спуск. Наприклад, існує m вхідних параметрів моделі, метод градієнтного спуску міні-серіями за одну ітерацію буде опрацьовувати $b < m$ значень. Тобто, навіть якщо навчальна вибірка є великою, вона буде оброблятися в невеликих тренувальних групах за один раз. Отже, даний метод є ефективним і для великих вибірок даних для навчання за рахунок меншої кількості ітерацій для оптимізації цільової функції [119].

4. Градієнтний спуск з моментом

Даний метод дозволяє пришвидшити спуск у певному напрямку, а також зменшує коливання, що дозволяє швидко наблизитись до локального мінімуму. Це можна реалізувати за рахунок додавання відповідної частини вектора оновлення з $t-1$ кроку до вектора поточного оновлення [119]:

$$\begin{aligned} v_t &= \gamma v_{t-1} + (1 - \beta) dW(x_t), \\ W &= W - \alpha v_t \end{aligned} \quad (3.21)$$

де v та dW – аналогічні прискоренням та швидкості;

α – швидкість навчання;

β – зазвичай зберігається при 0.9.

5. Метод Адама (англ. (Adaptive Moment Estimation))

Оптимізатор Адама є обчислювано ефективним, а також не має багато вимог до пам'яті. Саме тому, він є одним із найпопулярніших алгоритмів, які використовуються для оптимізації спуску градієнтів. Використовуючи оцінки першого та другого моментів, даний метод дозволяє розрахувати індивідуальну адаптивну швидкість навчання [119,121].

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t, \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \end{aligned} \quad (3.22)$$

де m_t та v_t – оцінки першого та другого моментів.

Параметри моделі оновлюються, використовуючи наступні закони [119]:

$$\begin{aligned} \widehat{m}_t &= \frac{m_t}{1 - \beta_1}, \\ \widehat{v}_t &= \frac{v_t}{1 - \beta_2}, \\ \theta_{t+1} &= \theta_t - \alpha \frac{m_t}{\sqrt{v_t + \varepsilon}}, \end{aligned} \quad (3.23)$$

Визначення та відбір даних, на яких нейронна мережа буде тренуватися - одним із найважливіших факторів під час її побудови. Метою побудови мереж у даній роботі є прогнозування подій, пов'язаних із розпізнаванням втручання та реагування системи кіберзахисту об'єкта на несанкціоновані втручання з метою забезпечення бази для прийняття рішень стосовно локалізації наслідків критичної ситуації чи загрози. Тобто, основна задача – це визначення індикаторів та

вхідних даних, які будуть використовуватися, а також у підборі необхідних навчальних даних.

Обсяг, тривалість, щоденна або погодинна зміна в кількості звернень до системи зазвичай використовуються як вхідні дані до моделі. Проте ANN може також отримувати як вхідні дані, наприклад, як ідентифікатор технічного засобу – джерела події, тип даних події, часова мітка події, IP-адреси джерела події та призначення події, статис події, сенсор, на якому виявлено подію, тощо [121].

Аналіз та відбір вхідних параметрів є першим та найважливішим кроком, який в подальшому впливає на якість навчання нейронної мережі. Наступним кроком є забезпечення вхідними даними так, щоб мережа мала можливість правильно навчитися без тренування.

Інший вид реалізації нейронної мережі – згорткова нейронна мережа, у якій взаємозв'язки між нейронами базуються на основі організації зорової кори тварин. Окремі нейрони такої мережі розміщені таким чином, що вони швидко реагують на перекриття або зміну області поля. Реалізація згорткової нейронної мережі найчастіше включає наступні кроки [119-126]:

1. Імплементация кожним фільтром згортки з метою формування карти збудження
2. Максимальна агрегація області інтересу, яка була отримана у попередньому кроці
3. Повторення 1-2 кроків до тих пір, поки не залишаться лише ознаки із необхідним рівнем
4. Застосування MLP (англ. Multi-layer Perceptron) для знаходження рішення конкретної задачі, використовуючи для вхідних параметрів дані із попередніх кроків.

Згорткова нейронна мережа характеризується використанням згорткових шарів для того, щоб визначити ознаки мережі. Це дає можливість проводити тренування нейронної мережі без складної попередньої обробки даних, так як більшість необхідних функції вивчаються вже безпосередньо під час навчання [18].

Рекурентна нейронна мережа (англ. Recurrent Neural Network RNN) – це штучна нейронна мережа, у якій нейрони один одному передають сигнали зворотного зв'язку [119, 121, 126]. У звичайній нейронній мережі, як правило, всі вхідні та вихідні дані є незалежними один від одного. Проте такий підхід не є ефективним для вирішення усіх завдань. Наприклад, якщо необхідно спрогнозувати наступне слово в реченні, то кращою буде та модель, у якій відомо, які слова були до нього. RNN виконують однакове завдання для кожного показника із вхідних даних, при цьому вихідні показники завжди залежать від попередніх результатів. Саме тому нейронна мережа і називається рекурентною. Варто зазначити, що в теорії RNN під час обчислень може брати до уваги й досить віддалену попередню інформацію, але, як показує практика, найчастіше обмежується кількома кроками назад [119, 121, 126].

Предсталені моделі штучного інтелекту використані для вирішення основної проблем дослідження – формування аналітичного інструментарію для прийняття рішень в умовах невизначеності та обмеженості в часі за різних сценаріїв розвитку подій та горизонту прогнозування. Система повинна бути спроектована таким чином, щоб ефективно визначати цілі та завдання, виходячи з специфіки критичної інфраструктури. Важливо забезпечити гнучкість в управлінні типами даних, включаючи інтеграцію з різними джерелами та забезпечення їхньої якості та актуальності. Для гарантії високого рівня безпеки і конфіденційності даних необхідно впровадити потужні механізми захисту інформації, такі як шифрування та контроль доступу. Надійність і доступність системи мають забезпечуватися через ефективні механізми резервування, відновлення та автоматичного відновлення після збоїв. Аналіз даних і прийняття рішень повинні базуватися на сучасних алгоритмах і моделях, зокрема методах аналітики та машинного навчання, що дозволить досягати високих результатів у виявленні та прогнозуванні можливих проблем. Користувацький інтерфейс має бути інтуїтивно зрозумілим та адаптованим до потреб різних категорій користувачів, що забезпечить зручність у використанні системи.

3.4 Застосування багатомодельного підходу

Процеси, характерні для розвитку більшості реальних складних систем та об'єктів, визначаються наявністю невизначеностей різних типів, спричинених як зовнішнім середовищем функціонування, так і їх складністю, нелінійністю, нестаціонарністю, високою динамічністю тощо. Все це приводить до того, що кожне рішення оцінюється за одним або декількома критеріями в рамках однієї або декількох моделей – тобто виникає потреба створення і застосування багатомодельного підходу до розв'язання складних задач.

Крім того, для виявлення закономірностей, що характеризують поведінку досліджуваних процесів, розроблено та широко застосовується сучасний математичний апарат аналізу часових рядів. Однак, у випадку наявності нелінійностей, виникають проблеми із практичним застосуванням теоретичних розробок. Значною мірою це пов'язано з тим, що такі часові ряди, крім складності структури ряду, часто містять пропуски, шумові складові, аномальні значення, є нестаціонарними, можуть містити локальні особливості, які можуть бути досить складними, неперіодичними ефектами. Все це значно ускладнює процес оцінювання структури і параметрів адекватної моделі для прогнозування. Тому на перший план виходять методи, які не завжди можуть мати теоретичне обґрунтування, а переважно засновані на практичному досвіді аналітика і розробника, а також вхідних умовах задач прогнозування і подальшого прийняття рішень з урахуванням отриманих результатів прогнозування.

Добре зарекомендували себе й потужні аналітичні платформи, в рамках яких реалізовані сучасні інформаційні технології, розраховані на використання досить значних обчислювальних потужностей та які вже містять широке коло можливих структур математичних моделей, потужний апарат інтелектуального аналізу структурованих та неструктурованих даних. [27, 49, 116, 127-129]. Це надає переваги у застосуванні адаптивного підходу при побудові моделей, а

також неформалізованого підходу при побудові допустимих ансамблів математичних моделей.

$$M = \{M_0(Y, I, P), M_E(X), M_{OE}, M_s, M_D(Q), M_{MO}, M_{ME}, M_U, A, M_H, M_{RS}, M_v\}, \quad (3.24)$$

де, $M_0(Y, I, P)$ - ідентифікуюча модель системи, Y - ендегенні змінні; I - вектор керованих змінних; P - вектор ресурсів; $M_E(X)$ - модель навколишнього середовища, X - екзогенні змінні, M_{OE} - модель взаємодії об'єкта та навколишнього середовища; M_s - модель створення синергетичного ефекту в системі; $M_D(Q)$ - модель поведінки системи, Q - збурюючі впливи; M_v - модель взаємодії з підсистемами інших рівнів; M_{MO} - модель зміни стану системи; M_{ME} - модель зміни стану навколишнього середовища, M_U - модель керуючої системи; A - правило вибору процесів зміни об'єкта, M_H - модель впливу особи, що приймає рішення на систему та результати дослідження; M_{RS} - модель системних ризиків.

З метою створення єдиної концептуальної основи для вирішення задач побудови інформаційних технологій прогнозування запропоновано системну методологію побудови моделей, розроблену на основі адаптивного підходу до моделювання з комбінованим використанням сценарного моделювання, регресійних та ймовірісно-статистичних моделей у формі мереж Байєса, багатовимірних розподілів та узагальнених лінійних моделей [9].

Її перевагами є комплексне застосування множини ймовірісно-статистичних методів виявлення та врахування невизначеностей (цифрова і оптимальна фільтрація, альтернативні методи оцінювання структури і параметрів математичних моделей), і забезпечення підвищення якості проміжних і остаточних результатів обробки даних, побудови математичних моделей і оцінювання прогнозів. Загальна схема застосування пропонованого підходу подана на рис. 3.2 []



Рисунок 3.2 – Загальна схема застосування пропонованого багатомодельного підходу

Як видно з рис. 3.2, пропонований підхід може бути застосований для прогнозування процесів, що мають місце у системах у соціально-економічній системі регіону. Перевагою розробки є те, що її легко «перелаштувати» на вирішення різних задач за різних вхідних умов: урахувати можливі варіанти розвитку ситуації, опрацювати короткі вибірки значної кількості різнорідних показників, експертних оцінок та можливих варіантів рішень. Важливе місце у розробленій методиці застосування багатомодельного підходу приділено саме всебічному дослідженню предметної області задачі, огляду варіантів розвитку подій, а також передбаченню їх наслідків на перспективу. Дана проблема вирішується шляхом розробки інформаційної технології, яка поєднує в собі процедури синтезу оптимального рішення, використання інтелектуального

аналізу даних, прогнозного моделювання, а також доступних інструментів підтримки прийняття рішень [7, 10-14, 18, 46, 47, 60].

Як видно з рис. 3.2, значна увага приділена забезпеченню системності використання методів аналізу та автоматизованого інтегрування різнотипної інформації, методів моделювання, прогнозування та багатокритеріального прийняття рішень. Для подолання невизначеностей запропоновано множину різних методів, залежно від типу невизначеності.

Очевидно, що за пропонованого підходу значна увага приділена методиці попереднього опрацювання значних обсягів різнорідних даних, отриманих із різних інформаційних джерел. Створена методика вирізняється тим, що при формуванні масивів вхідних даних їх інформативність, синхронність та коректність забезпечуються на етапі попередньої обробки даних, виконуваної для приведення їх до форми, яка забезпечить можливість коректного застосування методів оцінювання параметрів моделі та отримання їх статистично значущих оцінок. Крім того, на особливу увагу заслуговує проблема опрацювання пропусків в даних, оскільки у реальних часових рядах зазвичай наявні пропуски, спричинені різними факторами, і уникнути їх утворення не можливо, і, часто трапляється так, що вилучити такі дані з аналізу не можливо – це спотворить уяву про досліджуваний процес. Тому в даній методиці пропонується використовувати різні підходи до заповнення пропусків із різною кількістю пропущених значень. Опрацювання вхідних даних має бути системним та комплексним, і що головне, адаптовано до потреб користувача: поряд із заповненням пропусків даних, повинно виконуватись корегування значних імпульсних (екстремальних) значень, нормування вимірів у заданих межах, логарифмування великих значень та фільтрації шумових складових, виявлення та усунення мультиколінеарності. Передбачено також, що вхідні дані можуть бути використані у багатоваріантних розрахунках, коли розглядаються їх можливі зміни за різних сценаріїв розвитку досліджуваних процесів.

Впровадження технологій інтелектуального аналізу даних у інформаційну технологію прогнозування та підготовки варіантів рішень представлено на рис. 3.3.

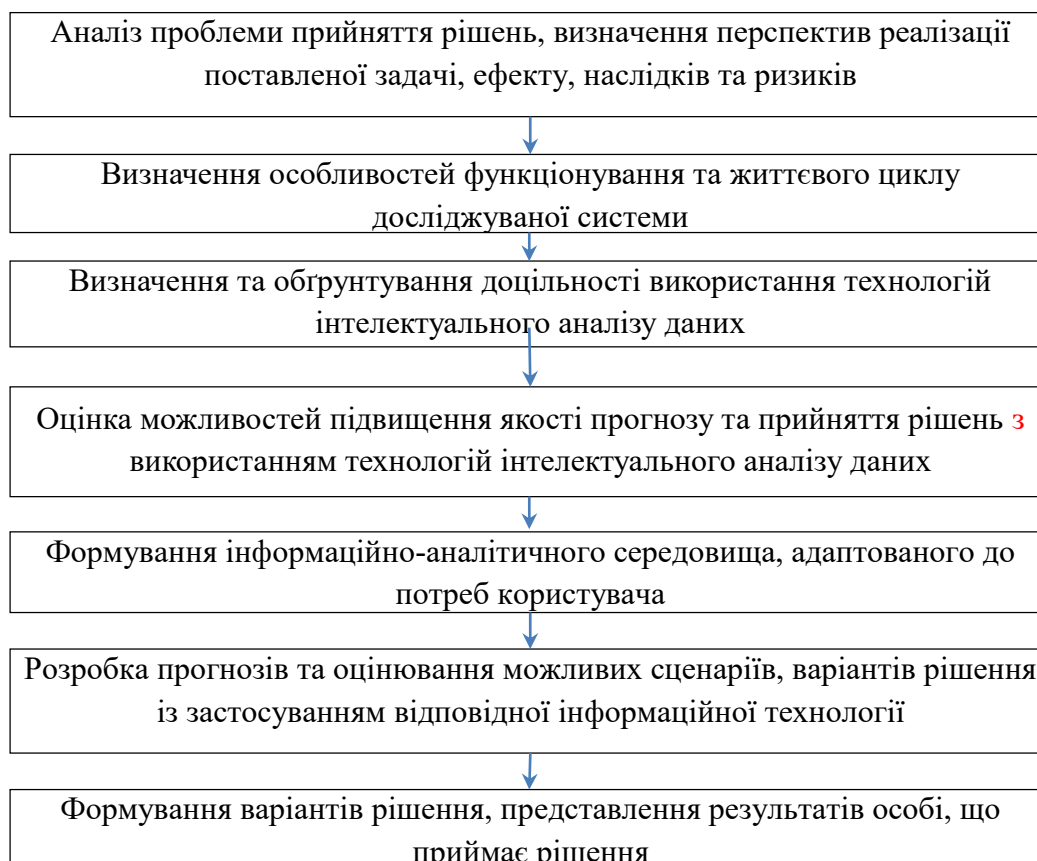


Рисунок 3.3 – Схема використання інтелектуального аналізу даних у інформаційній технології прогнозування для подальшого використання у інформаційній технології прийняття рішень

Запровадження у інформаційну технологію прогнозування сценарного підходу [116, 117], невід’ємною складовою якого є попередній аналіз предметної області, відбір альтернатив і формулювання цільових установок можливих варіантів розвитку подій, дозволяє розробляти прогнози навіть в умовах нечіткої, неповної інформації, оцінювати можливі ризики та невизначеності різних типів на основі процедури оцінювання і ймовірних варіантів розвитку ситуації, що забезпечує підвищення якості остаточних результатів.

Інформаційна технологія прогнозування, пропонована в роботі, поєднує в собі різноманітні методичні підходи до використання окремих методів та моделей а також їх комбінацій для отримання кінцевого результату – оптимального управлінського рішення – обґрунтованого прогнозними

розрахунками [59, 60, 66]. У таблиці 3.1 узагальнено основні методичні підходи, пропоновані до застосування на різних етапах.

Таблиця 3.1 – Методичні підходи, що реалізовані у інформаційній технології прогнозування

Етап	Методичні підходи
1. Збір інформації, формування інформаційної моделі досліджуваної предметної області	Емпіричні методи дослідження, метод статистичних спостережень, аналізу та синтезу, системного аналізу
2. Огляд стану та динаміки об'єкта дослідження, виявлення характерних ознак	Статистичний аналіз, data-mining, text-mining, факторний аналіз, багатовимірний аналіз даних, метод головних компонент, кореляційно-регресійний аналіз, типологічні та структурні групування, RFM-аналіз, кластерний аналіз, когнітивний аналіз
3. Причинно-наслідковий аналіз	Кореляційно-регресійний аналіз, ймовірнісне моделювання (мережі Байєса)
4. Розробка сценаріїв розвитку подій	Експертних оцінок, морфологічний аналіз, сценарне моделювання
5. Обґрунтування варіантів управлінських рішень	Data-mining, нейронні мережі, економетричне моделювання, SWOT-аналіз, когнітивне моделювання
6. Аналіз результатів	Метод експертних оцінок, метод Делфі, графічний

Як видно з таблиці 3.2, варіантів композицій застосування методичних підходів до аналізу та прогнозування розвитку є досить багато, всі вони опрацьовані.

Доповнюючи підхід, представлений в [7, 9, 16, 60, 64] у загальному сенсі, завдання багатокритеріального вибору кращого рішення на множині математичних моделей, можна представити так (3.26):

$$M = \{M_0(Y, I), M_E(X), M_{OE}, M_D(Q), M_{MO}, M_{ME}, M_U, A, M_H, M_{RS}, M_v, Z\}, \quad (3.26)$$

де $M_0(Y, I, P)$ – ідентифікуюча модель системи; Y – ендегенні змінні; I – вектор керованих змінних; $M_E(X)$ – модель навколишнього середовища; X – екзогенні змінні; M_{OE} – модель взаємодії об'єкта та навколишнього середовища; $M_D(Q)$ – модель поведінки системи; Q – збурюючі впливи; M_v – модель взаємодії з підсистемами інших рівнів; M_{MO} – модель зміни стану системи; M_{ME} – модель зміни стану навколишнього середовища; M_U – модель керуючої системи; A – правило вибору процесів зміни об'єкта; M_H – модель впливу особи, що приймає рішення, на систему та результати дослідження; M_{RS} – модель системних ризиків, Z – звороний зв'язок.

Проблема формування оптимальної методики прогнозного моделювання значною мірою зумовлена необхідністю одночасного подолання цілої низки невизначеностей різних типів: ситуаційної, статистичної, структурної, ймовірнісної тощо. Хоча, більшість існуючих інформаційних систем прогнозування та підтримки прийняття рішень мають відповідні засоби, однак, увага зосереджується на опрацюванні невизначеностей лише окремих типів. У даному дослідженні для вирішення цієї проблеми запропоновано інформаційну технологію застосування інтелектуального аналізу великих масивів структурованих та неструктурованих даних. Важливою перевагою пропонованого підходу є можливість розробки моделей для різних горизонтів прогнозування, адже від того, який горизонт прогнозування обрано, залежить не лише вибір типу моделі, а й підбір вхідних даних, зокрема, консолідація найбільш вагомих чинників. Крім того, необхідним є урахування тенденцій розвитку досліджуваних процесів, формально-математичний опис алгоритмів та методів генерування набору кандидатів-сценаріїв розвитку процесу (рис. 3.4) [64].

Етап	Характеристика етапу			
Завантаження вхідних даних	Завантаження часових ряди історичних даних			
Діагностика даних	Визначення аномалій та їх обробка (виключення, згладжування).			
Обробка пропусків	Заповнення пропусків			
Формування набору додаткових чинників	Додаються чинники, суттєві для аналізу досліджуваних процесів (дані, що надаються за бажанням користувача, довідкових дані, даних щодо екзогенних параметрів процесу, експертні оцінки, дані, що доповнюють відомості про аналізований процес)			
Строк	Дуже короткострокове	Короткострокове	Середньострокове	Довгострокове
Особливості побудови прогнозу	на кожную годину, враховуються цикли всередині доби.	на добу (тиждень) вперед, враховуючи цикли та закономірності в середині тижня.	Трендова складова, тренд та комбінований фактор з іншими регресорами, похідні показники	Враховуються додаткові чинники, користувача (на довгострокову перспективу)
Включення додаткових чинників	Додаткові регресори, що описують процес	Додаткові регресори, що описують процес.	Враховуються додаткові чинники, користувача	В залежності від предметної області
Горизонт прогнозу	від 1 до 24 годин	від доби до місяця	від 1 місяця до 3 років	більш 3 років
Сценарій	оптимістичний, реалістичний, песимістичний			
Побудова моделей-кандидатів	На кожному горизонті прогнозування будується модель за історичними даними та за даними користувача			
Етап 1	Одноступеневі моделі (експоненційні, регресійні, авторегресійні, узагальнені лінійні моделі)			
Етап 2	Двоступеневі моделі (регресійні та авторегресійні моделі. моделі із включенням трендової складової та урахуванням залишків (різниця між реальним та прогнозним значенням), що включаються до моделі у вигляді ковзного середнього, за умови, що між залишками та цільовою змінною є кореляція (автокореляція). Моделі класу експоненційного згладжування – Хольта, Тейла-Вейджа, Брауна, Вінтерса (з адитивною або мультиплікативною сезонною складовою), з урахуванням демпфуючого тренду та інші модифікації, Узагальнені лінійні моделі, нейронні мережі, ймовірнісні моделі та нечіткі методи.			
Вибір кращої моделі-кандидата	Вибір за результатами виконання етапу 1 та етапу 2, на основі: середньої абсолютної процентної похибки (MAPE), максимального MAPE, коефіцієнт детермінації (R^2), середньоквадратична похибка (RMSE).			
Побудова прогнозу	Використовуються кращі моделі-кандидати, вхідні дані: часові ряди досліджуваних процесів та додаткові чинники, що є значимими на даному горизонті прогнозування.			

Рисунок 3.4 – Методика застосування моделей різного типу та їх комбінацій залежно від горизонту прогнозування

Як видно з рис. 3.4, оцінювання якості побудованих сценаріїв, обґрунтування вибору найкращого та найімовірнішого з них виконано з використанням ймовірісно-статистичного моделювання. Для побудови сценаріїв можуть бути застосовані різні методики, що передбачають використання як кількісних, так і якісних показників, зокрема, соціально-демографічних, економічних, а в окремих випадках і суспільно-політичних. Для вирішення проблеми відсутності достатньо довгих часових рядів статистичних даних, розроблена методика формального опису причинно-наслідкових зв'язків. Особливістю методики є використання ймовірісно-статистичних методів для створення та застосування уніфікованих за структурою моделей в просторі станів. Для цього спочатку вирішується задача оцінювання елементів структури і параметрів моделі. Структура моделі оцінюється на підставі дослідження закономірностей протікання процесу, застосування статистичних тестів для перевірки наявності нелінійності, інтегрованості, гетероскедастичності, аналізу кореляційних функцій та візуального аналізу даних. При цьому вибирається кілька найбільш ймовірних структур моделей-кандидатів. Потім обчислюються оцінки параметрів моделей-кандидатів, обирається краща з них, використовуючи відповідні статистичні характеристики адекватності моделей. Ефективність застосування даного методу перевірено під час розв'язування практичних задач прогнозування регіонального розвитку ,. Також, для прогнозування фінансових та економічних процесів запропоновано метод на основі адаптивно-оптимізаційного підходу до моделювання та прогнозування з використанням регресійних, узагальнених лінійних та ймовірісно-статистичних моделей у формі мереж Байєса [6, 7, 45, 47, 49, 64-66, 116, 117, 128,129].

Застосування багатомодельного підходу виправдане і у випадках, коли дані, які описують досліджувані процеси, неповні або є сумніви щодо їх достовірності. Для розв'язання цієї задачі запропоновано багатомодельний підхід; а в основу цього методу покладено інтегроване використання теорії подібності процесів та ймовірісно-статистичного моделювання (рис. 3.5).



Рисунок 35 – Схема багатомодельного підходу із використанням теорії подібності процесів та ймовірісно-статистичного моделювання

Для оцінювання якості прогнозів, обчислених за побудованими моделями, запропоновано процедуру автоматизованого вибору кращої прогнозної моделі, у якій для вибору було використано інтегральний критерій якості, який включає 2-5 окремих статистичних критеріїв якості. Процедура автоматизації забезпечує можливість побудови, аналізу та вибору кращої із множини можливих моделей, кількість яких може сягати кількох сотень. Витрати часу залишаються при цьому цілком прийнятними для практичного використання пропонованої методики у автоматизованій системі, призначеній для прогнозного моделювання.

Отже, для розв'язання задач прогнозування необхідно створити такі інформаційні технології:

- аналізу та попередньої підготовки інформації;
- побудови моделей;
- прогнозування;
- опрацювання результатів прогнозування та представлення їх вигляді,

придатному для використання у процесі прийняття рішень.

Враховуючи те, що результати прогнозування створюють основу для прийняття ефективних рішень, проблема вибору оптимального рішення може

вирішуватися за використання багатомодельного підходу, який передбачає, що для обґрунтування вибору необхідно дослідити результати прогнозування, отримані із використанням множини різнорідних моделей, в тому числі і математичних.

Для створення інформаційної технології побудови моделей використано різні класи моделей, вибір яких здійснюється із застосуванням методології системного аналізу, ієрархічного підходу до створення та аналізу процесу моделювання, адаптації структури і параметрів моделей до особливостей досліджуваних процесів та на основі порівняння оцінок прогнозів за числовими критеріями їх якості.

3.5 Методика моніторингу результативності державної регіональної політики в умовах невизначеності

Зростаючі вимоги до сучасних інформаційних технологій з боку різних груп користувачів, і особливо з боку осіб, зацікавлених у підвищенні ефективності управління та зниженні ризику прийняття невірної управлінського рішення, спонукають до удосконалення існуючих методів та інформаційних технологій аналізу даних, пошуку нових рішень, що можуть бути швидко адаптовані до нових умов функціонування досліджуваних процесів.

У даній роботі пропонуються різні інформаційні технології в основу яких покладено поетапне розкриття невизначеностей різної природи, уточнення результатів на кожному етапі дослідження. Особливістю такого підходу є те, що особа, яка приймає рішення, має можливість отримати не лише «готове» рішення, а й брати участь у формуванні моделей, пропонувати власні сценарії та горизонти прогнозування для опрацювання з використанням фактичних даних. Інформаційна технологія в даному випадку повинна бути зручним і зрозумілим інструментом користувача, робити процес прогнозування більш прозорим та

обґрунтованим, приховуючи при цьому всю складність застосовуваного математичного та технічного інструментарію.

Застосування системного підходу у дослідженні складних соціально-еколого-економічних систем передбачає вивчення взаємодії елементів і підсистем із урахуванням динаміки зовнішнього середовища та внутрішніх зв'язків. У контексті горизонтального перетину регіональної системи, основними об'єктами дослідження виступають галузі національної економіки, фінансови сектор, ринок праці, зовнішньоекономічний сектор та інституційне середовище, тощо, а за вертикального переину – держава-регіон-район-громада..

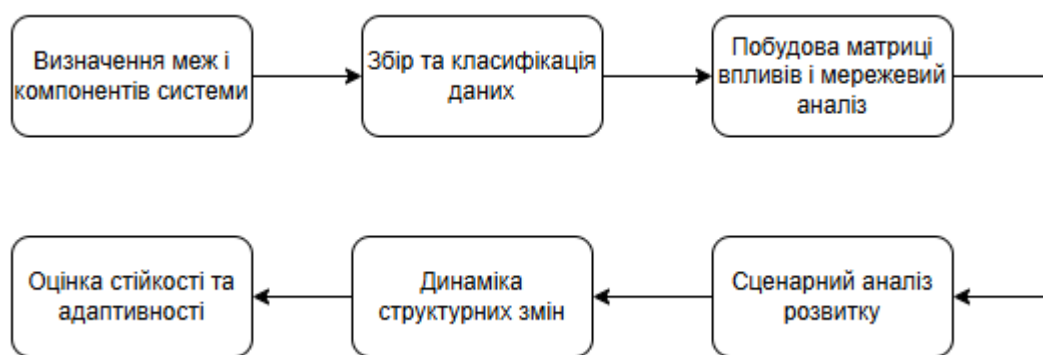


Рисунок 3.6 – Етапи дослідження складної соціально-економічної системи

Перший крок методики полягає в ідентифікації меж національної економіки України: територіальних (адміністративний устрій, зона конфлікту), функціональних (промисловість, сільське господарство, послуги, фінанси, ІТ-сектор, тощо) і часових (період 2015–2024 рр.) перетинів. У межах кожної підсистеми виокремлюються ключові елементи: в економіці — від великих корпорацій до дрібного бізнесу, у фінансовому секторі - від центрального банку до фінансових установ у містах та селах.

Територіальні межі соціально-економічної системи країни та її мезорівня визначені згідно адміністративно-територіального поділу відповідно до Конституції України. В якості мезорівня системи розглядають 24 області країни, Автономну Республіку Крим, Київ та Севастополь. На мезорівні, міста – обласні

центри (Харків, Одеса, Львів тощо) розглядаються з точки зору соціально-економічних центрів.

Методика збору даних включає використання порталу відкритих даних, даних Державної служби статистики України, Національного банку України, Міністерства фінансів України, Пенсійного фонду України, тощо з розбивкою за регіонами та інтеграцією даних міжнародних організацій, таких як World Bank, Європейський банк реконструкції та розвитку, Європейський парламент, тощо.

Функціональні межі можна окреслити використовуючи міжнародну класифікацію NACE Rev.2 або національну систему класифікації видів економічної діяльності (КВЕД).

Таблиця 3.2 – Приклад окреслення функціональних меж системи

Підсистема	Основні галузі (КВЕД)	Ключові компоненти
Промисловість	C (Обробна промисловість), D (Добувна промисловість)	Металургія (Metinvest), машинобудування («Южмаш»), хімія (ОПЗ)
Сільське господарство	A (Сільське, лісове господарство)	Агрохолдинги (Kernel, МХП), кооперативи, фермерські господарства
Послуги	G–N (Оптова/роздрібна торгівля, транспорт, готелі, освіта, медицина тощо)	Рітейл (АТБ, «Епіцентр»), логістичні оператори, приватні клініки, університети
Фінанси	K (Фінансова та страхова діяльність)	НБУ, комерційні банки (ПриватБанк, Ощадбанк), страхові компанії, FinTech- стартапи
ІТ-сектор	J62–J63 (Комп'ютерне програмування, консультування)	Експортери софту (SoftServe, EPAM), стартап-інкубатори, аутсорсингові компанії

Ключовими елементами в даній класифікації є:

- великі підприємства — формують до 30 % ВВП, мають значний мультиплікативний ефект.
- малий і середній бізнес — понад 90 % юридичних осіб, гнучкість, схильність до адаптації, важливі для зайнятості

- фінансові інститути — центральна ланка в розподілі капіталу, регулюванні грошово-кредитної політики
- інфраструктурні оператори енергетика (Укренерго, НАФТОГАЗ), транспорт (Укрзалізниця, аеропорти, тощо), зв'язок (Київстар, Vodafone, тощо).

Часові межі дослідження охоплюють період 2015–2024 рр., що дає змогу простежити результати реформ, помірне зростання в 2017–2019 рр., шоківий 2020 р. через пандемію COVID-19 із суворими карантинними обмеженнями, етап відновлення в 2021–початку 2022 рр., 2022 р. – початок повномасштабної війни, 2023 р. - період часткової стабілізації й відновлення у Обраний інтервал дозволяє виявити ключові тренди і структурні зрушення та включає два важливі кризові шоки — пандемію й війну — що дозволяє всебічно оцінити стійкість і адаптивність економіки.

Чітке визначення територіальних, функціональних та часових меж дозволяє сформулювати повноцінну модель національної економіки України як складної соціально-економічної системи. Це забезпечує основу для подальшої декомпозиції, побудови моделей взаємозв'язків та прогнозного сценарного аналізу.

Прикладом інформаційного дослідження показників соціально-економічного розвитку можна навести формування часових рядів даних на основі даних Державної служби статистики, Міністерства фінансів України, Національного банку України, тощо. Національного банку України (НБУ), Міжнародних організації (World Bank, IMF, OECD).

Дані Державної служби статистики України — це щорічні та щоквартальні звіти з показників ВВП, індексів споживчих цін, статистики зайнятості населення, рівня безробіття, обсягів промислового виробництва, сільського господарства, будівництва, торгівлі, транспорту тощо. Дані використовувалися в розрізі КВЕД Rev. 2 для класифікації за галузями та регіонами.

Національний банк України (НБУ) оприлюднює щомісячні статистичні бюлетені з показників грошово-кредитного ринку (монетарні агрегати М0–М3,

облікова ставка, обсяг кредитування економіки, тощо), курсу гривні, міжнародних резервів, платіжного балансу, прямих іноземних інвестицій.

Міжнародні організації (World Bank, IMF, OECD) формують статичні звіти за міжнародною методологією, застосовуються для верифікації даних та заповнення пропусків. В даному випадку використовувалися оцінки ВВП, інфляції та дефіциту бюджету.

Усього було опрацьовано понад 50 часових рядів із різною частотою (щомісячно, щоквартально, щорічно) за період 2015–2024 рр.

Аналіз починається з методики розрахунку таких показників як ВВП, сукупне споживання домогосподарств, державні видатки, валове нагромадження основного капіталу та чистий експорт. За даними Держстату, у 2017–2019 рр. середньорічний темп зростання ВВП становив 2,3–3,5 %, причому основним драйвером був внутрішній попит (споживання домогосподарств +2,5–3,0 %) та експорт агропродукції (+4,0 %). У 2020 р. через карантинні обмеження та скорочення інвестицій ВВП упав на 4 %. У 2022 р., за оцінками НБУ, економіка скоротилася майже на 30 % через руйнування виробничої та транспортної інфраструктури, а в 2023 р. розпочався етап відновлення з прогнозованим зростанням близько +4 %.

Основний показник, застосовуваний для лцінювання інфляції - індекс споживчих цін, що публікується із щомісячною періодичністю. Базова інфляція (без урахування адміністративно регульованих цін і сезонних продуктів) дозволила виділити трендову динаміку. У 2015 р. ІСЦ сягнув 13 % через девальвацію гривні, до 2019 р. знизився до 5 %, після чого у 2020 р. спостерігалось 4 % (ефект падіння попиту), а у 2022 р. — стрибок до 26 % внаслідок девальвації, логістичних перебоїв і зростання світових цін на енергоносії.

Показники безробіття вимірювались за національною методикою та методологією МОП: рівень безробіття як відношення кількості осіб, здатних і готових до роботи, до економічно активного населення. У 2015 р. безробіття

становило 9,5 %, до 2019 р. знизилося до 8 %, а у 2022 р. підскочило до 20 % через внутрішню міграцію, мобілізацію та руйнування підприємств.

Для оцінки ролі експорту та імпорту використовувалися щоквартальні дані НБУ за статтями платіжного балансу. Середня частка експорту становила близько 45 % ВВП, при цьому аграрна продукція забезпечувала 25 % експортних надходжень, металургія — 20 %. Імпорт енергоносіїв та обладнання коливався від 30 % до 35 % імпортової виручки.

Класифікація даних.

– Галузева: згідно з КВЕД Rev. 2, об'єднання в три секції — первинний сектор (A–B), вторинний (C–E) та третинний (F–S).

– Регіональна: розподіл показників за 24 областями, АР Крим та містами з особливим статусом; особлива увага до даних із зон бойових дій, де статистика має значні прогалини.

– Інституційна: за сектором інституційних одиниць (домогосподарства, нефінансові корпорації, фінансові установи, держава, закордонний сектор).

Попередня обробка даних передбачає такі кроки як.

– сезонна корекція, наприклад, методом X-12 ARIMA для щоквартальних та щомісячних рядів.

– інтерполяція пропущених значень у регіональних даних з використанням лінійного та поліноміального згладжування.

– верифікація — перехресна перевірка з даними Міжнародного валютного фонду та Світового банку, а також з аналітичними звітами EBRD.

Цей комплексний підхід до збору й класифікації даних забезпечує достовірну основу для побудови моделі системного аналізу економіки України, враховуючи як макроекономічні тренди, так і регіональні та галузеві особливості.

Крім чисельних показників, більшість методик, особливо, призначених для дослідження регіонів та громад, пропонують використання SWOT та PEST аналізу. Побудова матриці впливів методика, зарактерна більше для бізнесу,

однак, для поглибленішого аналізу можна запропонувати використання даного інструмента.

Побудова матриці впливів починається зі збору двох типів інформації: кількісних кореляційних зв'язків між галузями та макропоказниками (ВВП, інфляція, зайнятість, курс гривні тощо) і якісних оцінок групи з 12–15 експертів (представників Мінекономіки, НБУ, провідних галузевих асоціацій). Для кожного «галузь ↔ макропоказник» фіксується коефіцієнт кореляції (за даними за 2015–2023 рр.) й одночасно експерт дає суб'єктивну оцінку сили та знаку впливу за шкалою від –3 (дуже сильний негативний ефект) до +3 (дуже сильний позитивний ефект). Після цього кількісні й якісні дані нормуються та об'єднуються в єдину зважену оцінку.

Таблиця 3.3 – Матриця впливів

ІТ-сектор → ВВП	обчислений мультиплікатор показує, що кожні +1 % зростання обсягу експорту ІТ-послуг дають додаткові +2,5 % до ВВП загалом. Аналітики пов'язують це з високою доданою вартістю софту та ефектом технологічного поширення в суміжні галузі.
ІТ-сектор → зайнятість	зростання ІТ-робочих місць на 1 % призводить до +1,8 % в загальному рівні зайнятості, адже сектор активно розвиває аутсорсингові кластери по всій країні.
Енергетична залежність → інфляція	через імпорт енергоносіїв негативний шок у ціні на газ і нафту дає до –2,2 % до «чистої» інфляції, що відображає здорожчання виробництва і логістики.
Енергетична залежність → курс гривні	підвищення цін на енергоносії знижує стабільність курсу приблизно на –1,5 одиниці індексу валютних коливань, що вказує на тиск на платіжний баланс.
Аграрний експорт → ВВП	додатковий приріст експорту агропродукції на 1 % забезпечує +2,0 % до ВВП, оскільки Україна є нетто-експортером зернових.
Аграрний експорт → стабільність курсу	одночасно ця ж динаміка створює негативний ефект на валютну стабільність (–2,7), бо різкі коливання світових цін на зерно призводять до значних припливів/відпливів валюти.

У таблиці 1.3 зведено ключові показники впливу основних секторів економіки України на макропоказники: мультиплікатори ІТ-сектору для ВВП і зайнятості, коефіцієнти впливу енергетичної залежності на інфляцію та курс гривні, а також ефект аграрного експорту. Значення розраховані на основі поєднання статистичних кореляцій і експертних оцінок, де знак вказує напрямок впливу, а числова величина — його силу.

На базі цієї матриці будується мережевий граф: вузли — це сектори й макропоказники, ребра — впливи зі значенням абсолютної сили понад 1,0. Далі застосовують стандартні метрики мережевого аналізу. Degree centrality (ступінь вузла) показує, які елементи мають найбільшу кількість прямих зв'язків.

- Betweenness centrality виявляє вузли, що найчастіше лежать на «коротких шляхах» між іншими парами — тобто виступають «містками» інформації.
- Closeness centrality вимірює середню відстань від вузла до всіх інших, що характеризує оперативність впливу.

За цими показниками - центральними вузлами мережі виявляються фінансова сфера (через прямі зв'язки з усіма макропоказниками: кредитування, облікова ставка, міжнародні резерви) та логістика (через мультиплікативний ефект на експорт, внутрішній ринок і енергетику). Їхній стан визначає загальну стійкість системи: навіть помірні коливання в цих секторах миттєво поширюються по всій мережі, що підкреслює необхідність пріоритетного моніторингу та управління ними.

У сценарному аналізі кожен з трьох сценаріїв спирається на чіткі допущення щодо зовнішнього середовища, внутрішніх реформ та темпів відбудови. У таблиці 1.4 наведено докладні характеристики кожного сценарію, їхні ключові допущення, макроекономічні траєкторії та секторальні ефекти.

Базовий сценарій моделює розвиток економіки за умов локалізації бойових дій у вузькому коридорі, при повільному відновленні інфраструктури та помірному темпі реформ. За збереження нинішнього рівня міжнародної допомоги й стабільної світової кон'юнктури очікується щорічне зростання ВВП

на 3 %, інфляція 8–10 %, поступове зниження безробіття до 12 % і курс гривні в діапазоні 40–45 грн/\$.

Таблиця 3.4 - Базовий сценарій розвитку

Допущення:	– Війну вдалося локалізувати, але активні бойові дії тривають у вузькому коридорі — основна інфраструктура відновлюється повільно. – Реформи тривають у нинішньому темпі: нові закони ухвалюються, але їхня імплементація гальмується бюрократією. – Міжнародна допомога залишається на рівні поточних траншів, приток прямих іноземних інвестицій помірний. – Цінова кон'юнктура на світових ринках помірно сприятлива (ціни на зерно й метал залишаються близько до середньорічних 2023 р. рівнів).
Макроекономічні показники (щорічно, 2024–2027):	– ВВП: +3 % – Інфляція: 8–10 % – Безробіття: з поточних ~15 % поступово до 12 % – Бюджетне сальдо: дефіцит 3–4 % ВВП – Курс гривні: коливається в діапазоні 40–45 грн/дол.
Секторальні ефекти	– Промисловість: відновлення виробництва на 60 % від довоєнного рівня, зростання +2 % щорічно. – Аграрний сектор: врожайність стабільна, експорт +1,5 % щорічно. – ІТ-сектор: зростання +8 % на рік, завдяки відновленню офісів і поверненню частини фахівців. – Будівництво та інфраструктура: реконструкція 30 % зруйнованих об'єктів до 2027 р.

Оптимістичний сценарій передбачає згасання бойових дій, статус кандидата в ЄС і реформи, що залучають інвестиції. ВВП зростає на 5–6 %, інфляція 5–7 %, безробіття падає до <10 %, курс — 35–38 грн/\$. Промисловість

+6 %, агроекспорт +5 %, ІТ +15 %, 40 % енергоспоживання з ВДЕ, відновлено 70 % інфраструктури.

Таблиця 3 5 - Оптимістичний сценарій

Допущення:	– Війна суттєво згасає протягом 2024 р., великий обсяг територій звільнених, що дає поштовх відбудові. – Україна отримує статус кандидата в ЄС і починає переговори про членство, що стимулює приплив інвестицій. – Реалізовано ключові антикорупційні та судові реформи, спрощено бізнес-регулювання. – Світові ціни на сировину зростають на 10–15 %, що підтримує експортний дохід.
Макроекономічні показники (щорічно, 2024–2027)	– ВВП: +5–6 % – Інфляція: 5–7 % – Безробіття: з поточних ~15 % до <10 % – Бюджетне сальдо: близько до нуля, за рахунок зростання податкових надходжень – Курс гривні: зміцнення до 35–38 грн/дол.
Секторальні ефекти:	– Промисловість: реальне зростання +6 % за рахунок модернізації заводів та відкриття нових виробничих ліній. – Аграрний сектор: експорт +5 % щорічно, активне залучення іноземних інвесторів у зберігання та переробку. – ІТ-сектор: подвійне зростання (+15 % щорічно) завдяки притоку стартап-капіталу та європейським проектам. – Енергетика: до 40 % споживання покривають відновлювані джерела, що знижує залежність від імпорту. – Інфраструктура: відбудовано до 70 % критичних об'єктів, запуснені ключові логістичні коридори.

Песимістичний сценарій передбачає затяжну війну, зупинку реформ і скорочення зовнішньої підтримки. ВВП падає на –5 % (далі –1 % щорічно), інфляція перевищує 20 %, безробіття – понад 25 %, дефіцит бюджету – 7–9 % ВВП, курс гривні – 50–55 грн/\$. Промисловість скорочується на –10 % щорічно,

агроекспорт падає на 3–5 %, IT-сектор стагнує (–2 %), енергетика страждає від відключень, а інфраструктуру відновлено менш ніж на 20 %.

Таблиця 3.6 - Песимістичний сценарій

Допущення:	– Війна затягується ще на 2–3 роки, активні бойові дії тривають у кількох регіонах. – Реформи зупиняються через політичну нестабільність, корупція посилюється. – Міжнародна фінансова підтримка скорочується, інвестори відкликають капітал. – Світові ціни на енергоносії різко коливаються, що посилює шоки в економіці.
Макроекономічні показники (2024–2027):	– ВВП: –5 % (одноразовий шок) і –1 % щорічно далі – Інфляція: >20 % (стагфляція) – Безробіття: >25 % – Бюджетне сальдо: дефіцит 7–9 % ВВП – Курс гривні: девальвація до 50–55 грн/дол.
Секторальні ефекти:	– Промисловість: падіння обсягів –10 % щорічно, масове закриття підприємств. – Аграрний сектор: зниження експорту на 3–5 % через руйнування логістики та аграрних потужностей. – IT-сектор: стагнація або незначне падіння (–2 %), масова еміграція фахівців. – Енергетика: часті відключення, залежність від дорогого імпорту. – Інфраструктура: відбудовано <20 % зруйнованих об'єктів, транспортні коридори не відновлені.

Критичними драйверами ризику є безпека енергопостачання: Будь-які перебої з імпортом газу/нафти чи затримки в розвитку ВДЕ одразу підвищують інфляцію й тиск на бюджет. Темпи євроінтеграції: Прискорена імплементація євростандартів відкриває доступ до фондів ЄС і ринку, уповільнення — відлякує інвесторів. Відновлення інфраструктури: Швидкість ремонту доріг, портів, енергомереж визначає можливості для відновлення торгівлі й промисловості.

Усі три драйвери діють синергетично: наприклад, покращення енергобезпеки прискорює індустріальний розвиток, а це, у свою чергу, стимулює інвестиції й підтримує курс гривні.

У 2015 році частка сектору послуг у ВВП України становила близько 60 %, а вже до 2021 р. зросла до 68 %. Такий приріст зумовлено як нарощенням внутрішнього споживання, так і розширенням експортно-орієнтованих сервісів. Зокрема, оптова й роздрібна торгівля збільшили свою вагу з 14 % до 17 % ВВП, логістика — з 8 % до 11 %, фінансові послуги — з 6 % до 8 %, сектор нерухомості й оренди — з 4 % до 6 %. Туризм приніс країні близько \$5 млрд у 2019 р., але через COVID-19 обсяги скоротилися до \$1,5 млрд у 2020 р.; уже в 2021–2022 рр. почалося відновлення внутрішнього турпотоку й онлайн-послуг (дистанційна освіта, телемедицина).

ІТ-галузь демонструє найдинамічніший ріст: експорт ІТ-послуг виріс із \$2 млрд у 2015 р. до \$7 млрд у 2021 р., упав до \$5 млрд у 2022 р. через масову мобілізацію фахівців і перебої з інфраструктурою, але вже в 2023 р. перевищив довоєнний рівень. Кількість зайнятих у секторі зросла з близько 100 тис. осіб у 2015 р. до 300 тис. у 2021 р., середня місячна зарплата піднялася з \$1 000 до \$2 000. Основні центри розвитку: Київ (40 % усіх компаній), Львів (15 %), Харків (10 %), Одеса (8 %). Створення ІТ-хабів та стартап-інкубаторів дало поштовх розвитку продуктових компаній поряд із класичним аутсорсингом.

Водночас важка промисловість зазнала суттєвого скорочення: частка у ВВП впала з 15 % у 2015 р. до 10 % у 2023 р. Основні причини — дефіцит енергоресурсів, руйнування металургійних та машинобудівних заводів на сході країни, а також перехід частини виробництв у режим простою. Виробництво сталі знизилося з 32 млн тонн у 2015 р. до 25 млн тонн у 2021 р., машинобудування скоротило обсяги на 20 %, хімічна промисловість — на 15 %. Металургійний експорт упав із \$16 млрд до \$10 млрд за той самий період. У 2022 р. через воєнні дії виробничі потужності Донбасу були виведені з експлуатації на 50 %, що ще більше посилило спад, а в 2023 р. почалося повільне

відновлення завдяки перезапуску ключових підприємств і програмі держпідтримки енергоефективності.

Оцінка стійкості та адаптивності національної економіки України можн бути здійснена на основі композитного індексу економічної стійкості (ІЕС), що поєднує три ключові компоненти:

Резерви НБУ / імпортні платежі (RR): Відношення міжнародних резервів до середньомісячного обсягу імпортних платежів. У 2019 р. цей показник становив близько 8 місяців ($RR_{19} \approx 8$), у 2022 р. – 3 місяці ($RR_{22} \approx 3$), у 2023 р. – 5 місяців ($RR_{23} \approx 5$).

Державний борг / ВВП (DR): Співвідношення державного боргу до ВВП. У 2019 р. $DR_{19} \approx 60\%$, у 2022 р. $DR_{22} \approx 70\%$, у 2023 р. $DR_{23} \approx 65\%$.

Ліквідність банківського сектору (LR): Коефіцієнт покриття (LCR) — відношення високоякісних ліквідних активів до короткострокових зобов'язань. У 2019 р. $LCR_{19} \approx 120\%$, у 2022 р. $LCR_{22} \approx 90\%$, у 2023 р. $LCR_{23} \approx 110\%$.

Для кожного компонента розраховано нормалізовану оцінку $[0; 1]$

$$x' = \begin{cases} \frac{x - x_{\min}}{x_{\max} - x_{\min}}, & \text{якщо більші значення — кращі (RR, LR);} \\ 1 - \frac{x - x_{\min}}{x_{\max} - x_{\min}}, & \text{якщо менші значення — кращі (DR).} \end{cases}$$

Де $x_{\min}$ і $x_{\max}$ — відповідно мінімальне й максимальне значення показника за 2015–2023 рр. Індекс економічної стійкості обчислюється як середньозважене (ваги 1/3) нормалізованих RR' , DR' та LR' :

$$ІЕС = \frac{1}{3}(RR' + DR' + LR').$$

Отже, тримано:

- $ІЕС_{19} \approx 0.75$ (висока стійкість до зовнішніх шоків завдяки значним резервам і здоровому банківському сектору).
- $ІЕС_{22} \approx 0.42$ (різке падіння через війну: витрати резервів на імпорт військового обладнання, зростання боргу, зниження ліквідності банків).

— $IES_{23} \approx 0.55$ (часткове відновлення: поповнення резервів за рахунок міжнародної допомоги, рефінансування боргу, зростання депозитної бази).

Як можна побачити з наведеного, сценарний аналіз одним із важливих засобів дослідження розвитку соціально-економічних систем, досягнення цілбових настанов позитивного сценарію. Сценарій засвідчили, що ключовими факторами, які сприяли забезпеченню економічної стабільності були гнучкість фінансової політики НБУ, міжнародна допомога та кредитні програми (понад \$20 млрд макрофінансової підтримки МВФ, ЄС, Світовий банк), диверсифікація експорту (орієнтація на ринки Азії, Латинської Америки), що зменшило вразливість до цінових коливань у традиційних регіонах.

Таким чином, застосування процедур системного аналізу до економіки України дозволяє виявити та ранжувати ключові компоненти й їх взаємозв'язки. Побудувати прогностичні сценарії розвитку з урахуванням ризиків. Оцінити стійкість та адаптивність національної економіки й сформулювати рекомендації для підвищення її ефективності в умовах невизначеності.

Висновки до розділу 3

1. У розділі представлено результати аналізу інформаційно-аналітичне забезпечення процесів, характерних для складних соціально-економічних систем – регіону та громади.
2. Встановлено, що найбільш проблемним питанням у процесі розв'язання задач прийняття рішень в управлінні регіональним розвитком залишається організація збору та оброблення даних, необхідних для аналізу стану та динаміки досліджуваних систем для подальшого їх використання для побудови моделей та прогнозів, які, як правило, супроводжується невизначеностями різного типу.
3. Визначено, що збір необхідних даних ускладнюється не тільки ситуацією військового конфлікту, а й змінами у методології статистичних досліджень, реформуванням системи місцевого

самоврядування, змінами у адміністративно-територіальному устрої. Зокрема, це невизначеності структурні, статистичні й параметричні.

4. Встановлено, що окремою проблемою є забезпечення якісної попередньої обробки даних, заповнення пропусків та відстеження аномальних значень.
5. Запропоновано методикау попередньої обробки даних, обов'язковою складової, якої є моделі та методи оброблення пропусків.
6. Визначено, що проблемним є й питання забезпечення своєчасного виявлення проблем реалізації заходів державної регіональної політики, тому, в роботі запропоновано інформаційну технологію для її моніторингу, яка на відміну від наявних рішень, поєднала математичні моделі, методи інтелектуального аналізу даних і штучного інтелекту з рекомендаційною підсистемою для формування сценаріїв і варіантів управлінських рішень в умовах невизначеності, що дає змогу своєчасно виявляти відхилення та пропонувати оптимальні коригуючі заходи.
7. На базі процедур системного аналізу, розроблена методика, яка дозволяє не лише об'єктивно оцінити поточний стан системи, а й запропонувати варіанти її розвитку, виявляти критичні елементи та забезпечувати гнучкість управління в умовах невизначеності.
8. Розроблено підхід, реалізований у даній інформаційній технології враховує особливості збору статистичних даних та періодичність виконання аналітичних процедур у відповідних органах управління і дозволяє використовуючи дані, що описують різні аспекти соціально-економічної ситуації дослідити результативність заходів державної регіональної політики.

РОЗДІЛ 4

МОДЕЛІ ТА МЕТОДИ ПРОГНОЗУВАННЯ РЕГІОНАЛЬНОГО РОЗВИТКУ

4.1 Методичні засади прогнозування процесів регіонального розвитку

Реформа децентралізації змінила концептуальні підходи до управління регіональним розвитком – територіальні громади отримали нові можливості та повноваження, а значна їх кількість об'єднались, утворивши нові громади – кардинально змінилась структура регіональної системи. Методичні підходи до розроблення стратегій, планів, програм регіонального розвитку, основою яких є виключно ретроспективні дані, за таких умов не забезпечують високу якість прогнозів. Потребує удосконалення як організаційна складова аналітичної роботи, так і методичні підходи до побудови прогнозів. Для удосконалення організації аналітичної роботи в громадах та на рівні регіону пропонується створення єдиного інформаційне середовище, яке має інтероперабельну мережецентричну архітектуру, комплекс інтелектуальних технологій опрацювання даних, що дозволяє забезпечити узгодженість взаємодії «громада-регіон», забезпечить зворотній зв'язок у системі управління регіональним розвитком, що дозволить підвищити обґрунтованість управлінських рішень. Для удосконалення методичних підходів до прогнозування регіонального розвитку запропоновано єдину формалізовану методологію прогнозування, яка охоплює повний цикл: від збору та інтеграції різномірних структурованих і неструктурованих даних до формування проекту рішення та базується на системному підході з використанням багатомодельного і багатокритеріального аналізу (рис. 4.1).

Основу пропонованої інформаційної технології становлять сучасні засоби збору та оброблення даних, методи інтелектуального аналізу даних, математичні моделі, штучний інтелект, методи підтримки прийняття рішень. Для підвищення якості аналізу стану досліджуваної системи у процесі розроблення стратегій,

програм, проєктів, тощо на рівні регіону та громади запропоновано використовувати систему збору та аналізу даних, яка включає використання SWOT, PEST, побудову матриць впливу, когнітивний та сценарний аналіз, а у для підготовки даних для побудови моделей – методи заповнення пропусків, інших методи аналізу структурованих та неструктурованих даних, отриманих з різних джерел, в тому числі й даних з мережі Інтернет

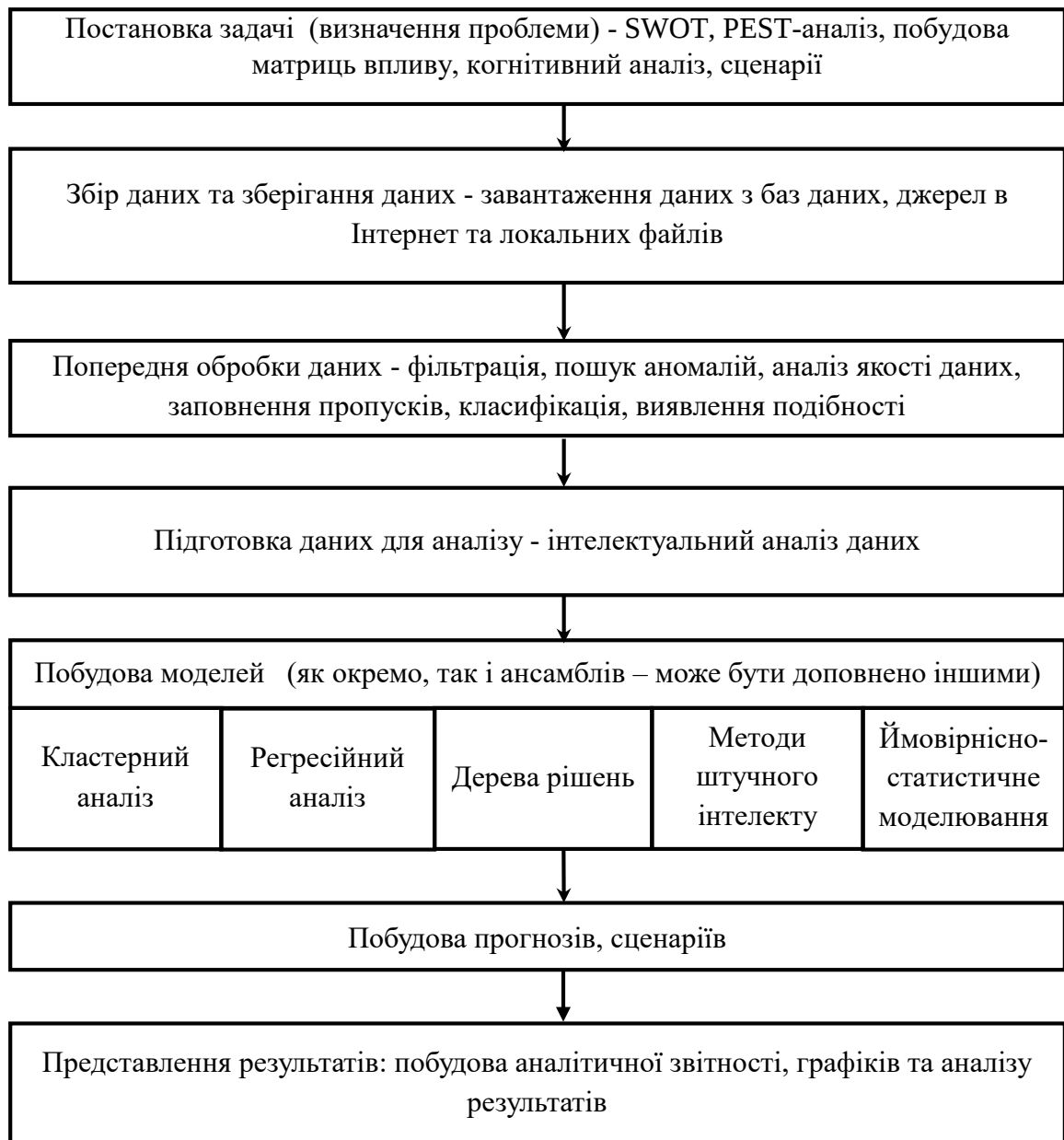


Рисунок 4.1 – Узагальнена схема побудови прогнозів

Передбачено, що для розв’язання окремих задач управління регіональним розвитком можуть бути використані сценарії, які містять встановлені обмеження щодо значень індикаторів чи показників соціально-економічного розвитку, тоді,

методика прогнозування буде використана таким чином, що сценарій буде завантажений з бази знань або завантажений користувачем.



Рисунок 4.2 – Етапи підготовки даних до використання у СППР

Пропонований підхід передбачає ієрархічний аналіз процесів моделювання та прогнозування, урахування невизначеностей структурного, параметричного та статистичного характеру, адаптування структури і параметрів моделей до змін у процесах та застосування альтернативних методів оцінювання параметрів моделей з метою пошуку кращих оцінок прогнозів за допомогою множини числових критеріїв їх якості. Значна увага приділена питанню прогнозування динаміки розвитку процесів. Прогнозування стосується таких складових процесу: детермінованого тренду як показника довгострокових змін соціально-економічного процесу; нерегулярного тренду як показника коротко- та середньострокових змін – випадкова складова тренду; прогнозування коливань або відхилень, що накладаються на тренд – це короткострокові зміни; оцінювання прогнозів сезонних ефектів; аналіз та прогнозування приростів або швидкості змін процесу в часі, які визначаються першими різницями; оцінювання прогнозу дисперсії або її стандартного відхилення (волатильності як міри розсіювання); прогнозування якісних змінних за допомогою нечітких множин, мереж Байєса; оцінювання комбінацій оцінок прогнозів, що обчислюються на множині індивідуальних оцінок, отриманих за допомогою вибраних методів моделювання і прогнозування досліджуваних процесів.

Для підвищення якості прогнозів, удосконалено метод адаптивного прогнозування процесів довільної природи для розв’язання задач прийняття рішень стосовно управління регіональним розвитком за рахунок використання у якості вхідних даних характеристики досліджуваних процесів, представлені у вигляді текстових повідомлень у мережі Інтернет.

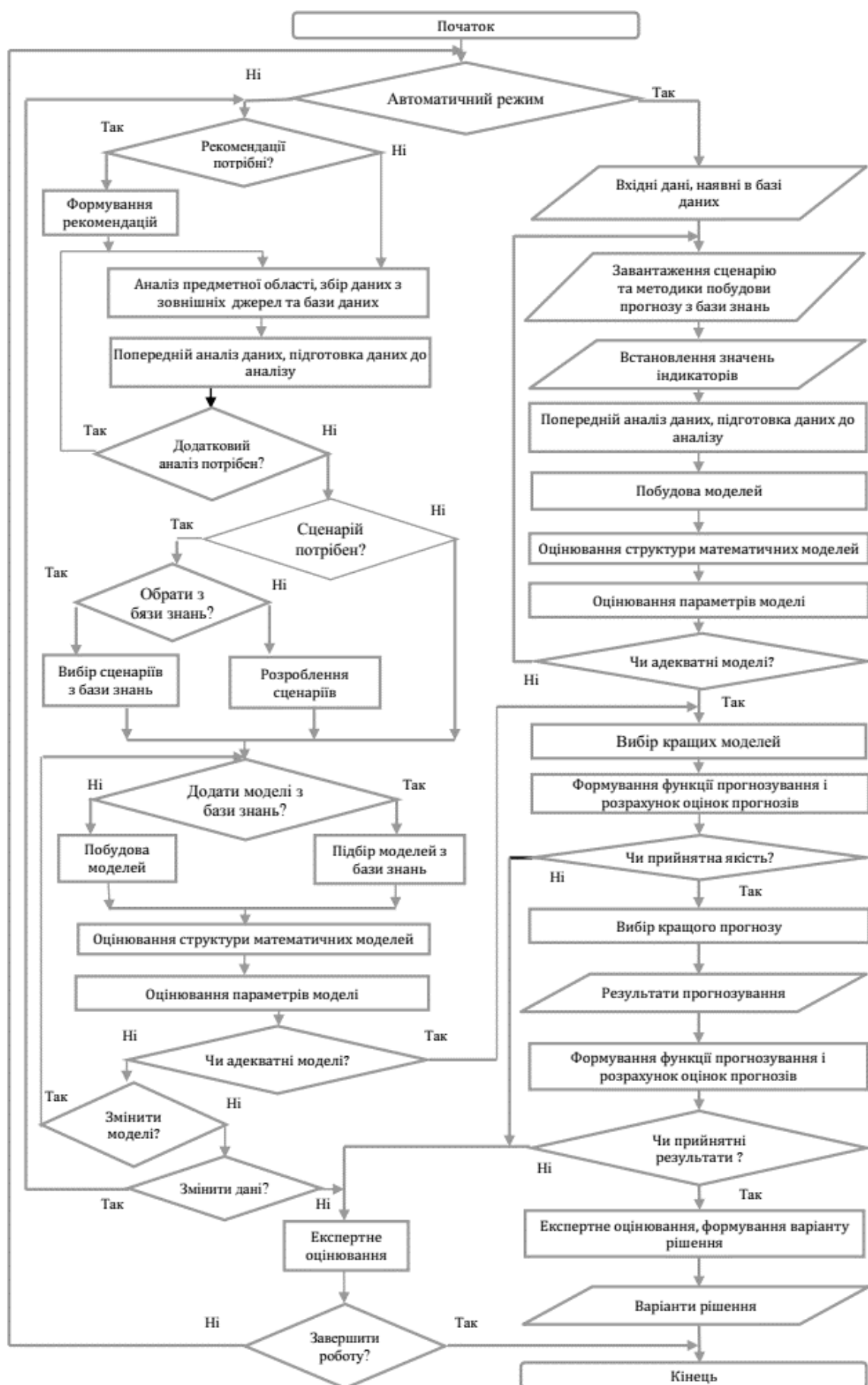


Рисунок 4.3 – Схема пропонованої методики адаптивного прогнозування

Пропонована методика вирізняється адаптивністю, масштабованістю, урахуванням невизначеностей різних типів та можливістю дослідження економічних, екологічних, соціальних чинників, що впливають на розвиток громад та регіонів, та особливостей управлінських та інформаційних зв'язків на різних рівнях управління системи «держава-регіон-громада», як в умовах війни, так і у період повоєнного відновлення.

4.2 Застосування методів прогнозування на основі часових рядів

Методи прогнозування на основі використання часових рядів - одні з найбільш прийнятних до використання у комп'ютерному моделюванні. Адже існує значна кількість методів прогнозування лінійних та нелінійних процесів з використанням даних у формі часових рядів [138, 156]. Найбільш поширеними серед них є метод групового врахування аргументів (МГВА), авторегресія (АР), та її різновиди ARMAX, з дробовою інтегрованістю ARFIMA, АРКС, авторегресія з інтегрованим ковзним середнім (АРІКС), нелінійна множинна регресія, квантильна регресія, регресійні дерева, нейромережі, байєсівські мережі, нечіткі множини, нечіткі нейромережі та інші. Відносно „універсальними” методами моделювання та прогнозування є МГВА і нечіткі нейромережі.

Однак, одного, навіть досить універсального методу, недостатньо для досягнення повноти аналізу процесу. Так, наприклад коректний аналіз гетероскедастичних процесів вимагає застосування моделей спеціальної структури для описування умовної дисперсії, що не забезпечує МГВА. Кожний метод має також свої недоліки і переваги стосовно обчислювальних витрат та характеристик точності оцінок прогнозів. Висока точність прогнозу за допомогою МГВА або нейромережі іноді досягається за рахунок високих обчислювальних витрат і моделей складної структури.

Це особливо стосується використання моделей вказаного типу у задачах дослідження розвитку соціально-економічних систем під впливом різних груп чинників, коли модель необхідна і для побудови прогнозу, і для використання у процесі прийняття рішення.

Суттєву перевагу щодо обчислювальних витрат можна досягти у такому випадку за допомогою набагато простіших моделей, наприклад АРКС (або АРІКС), перевагами яких є простота структури та можливості їх оперативної адаптації до характеристик процесу в реальному часі. Цьому сприяє також наявність множини відносно простих методів рекурсивного оцінювання параметрів моделей.

Крім того, представлена загальна методика обробки інформації при розв'язанні задач формування інформаційної бази для дослідження соціально-економічних процесів, в тому числі ідентифікацію та опрацювання невизначеностей різних типів при формуванні часових рядів вхідних даних, передбачає використання комплексу методів та моделей для моделювання та прогнозування та шляхи їх реалізації в рамках відповідної інформаційної технології.

Єдиного систематизованого підходу до вибору структури математичних моделей та методів прогнозування, а також рекомендацій стосовно їх застосування на даний час немає. Загалом, методи прогнозування можна розділити на три класи:

- методи прогнозування на основі суджень, тобто, прогнозування, на суб'єктивних судженнях (оцінках), інтуїції, поглиблених знаннях конкретної області та іншій інформації, що має відношення до прогнозованого процесу – так зване передбачення.
- методи прогнозування на основі використання часового ряду однієї змінної, тобто, на основі авторегресії з ковзним середнім (АРКС) та АРКС плюс модель тренду.

- методи прогнозування на основі використання часових рядів декількох змінних.

В останньому випадку ендогенна змінна, що прогнозується, залежить від декількох регресорів або екзогенних змінних у правій частині рівняння. Очевидно, що в загальному випадку метод прогнозування може поєднувати у собі кілька наведених методів. Застосування такого підходу забезпечує отримання високої якості прогнозів та прийнятих рішень, які на них ґрунтуються.

Схема процесу аналізу даних та прогнозування на основі часових рядів представлена на рис. 4.4.

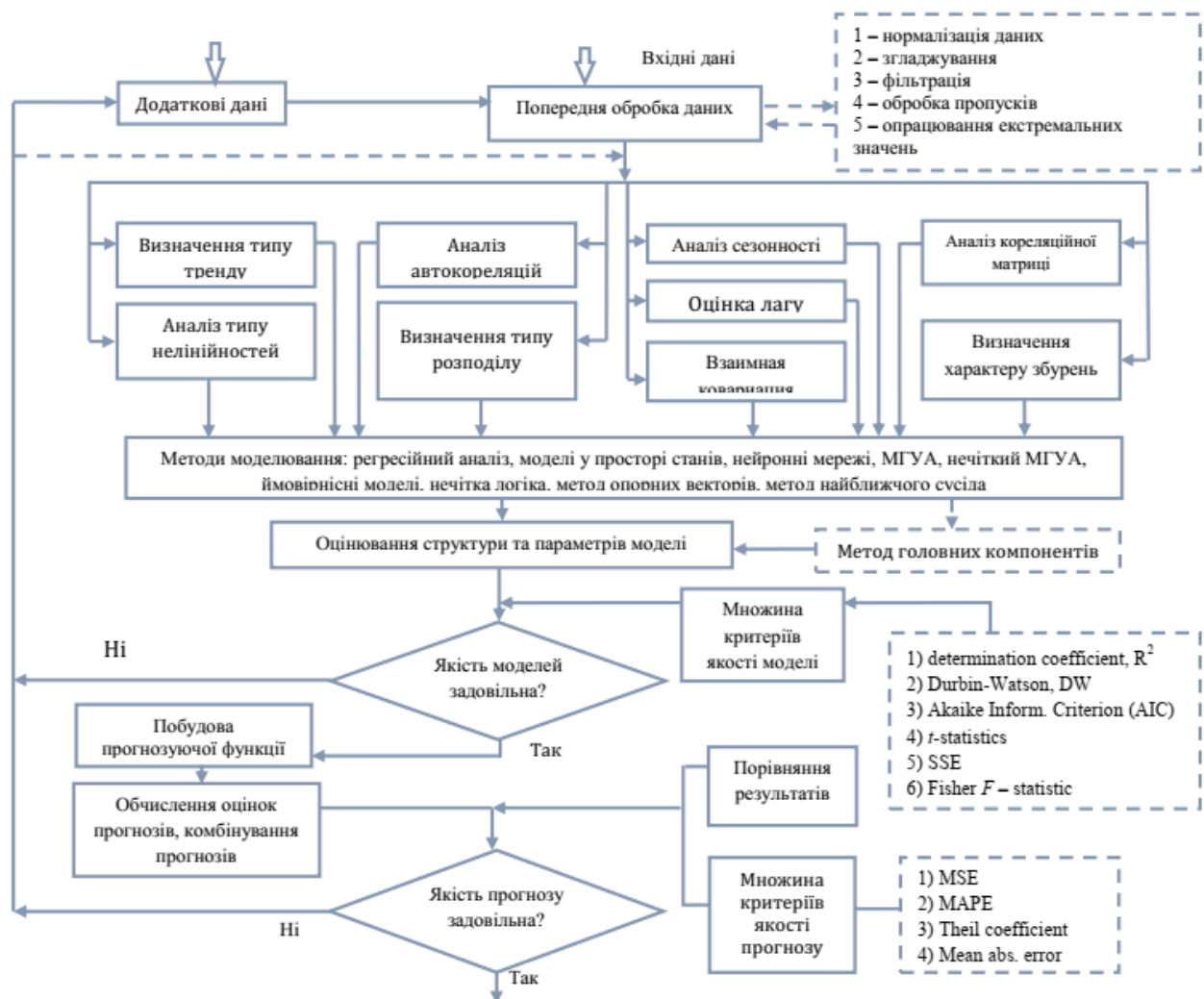


Рисунок 4.4 – Схема прогнозування соціально-економічних процесів на основі часових рядів (фрагмент)

Методика прогнозування розвитку соціально-економічних систем на основі часових рядів, призначена для комп'ютерного моделювання та прогнозування представлена в таблиці 4.1 [138, 156].

Таблиця 4.1 – Послідовність процедур прогнозування

Крок	Процедура	Склад процедури
1	Попередня обробка та аналіз даних	Заповнення пропусків; згладжування екстремальних значень; логарифмування; нормування в діапазоні від -1 до $+1$; диференціювання (можливості обчислення перших різниць та різниць вищих порядків); пряме і зворотне перетворення Фур'є; цифрова фільтрація; Перехід на крок 2.
2	Перевірка наявності нелінійностей	Перевірка наявності нелінійностей визначається за допомогою тесту Фішера або кореляційних функцій вищих порядків. Якщо процес містить вивода, то переходимо до кроку 3, інакше до кроку 4.
3	Визначення порядку вивода	Побудова моделі за МГВА і лінійна апроксимація процесу з метою її порівняння з нелінійною моделлю. Перехід на крок 9.
4	Перевірка процесу на стаціонарність	Перевірка процесу на стаціонарність за допомогою тесту Дікі-Фуллера. Якщо процес стаціонарний, переходимо до кроку 9, інакше до кроку 5.
5	Перевірки наявності гетероскедастичності	Для перевірки наявності гетероскедастичності застосовую такі тести: Уайта, Годфрі, Голдфельда-Квандта. Наявність тренду визначається за допомогою тесту Дікі-Фуллера. Якщо процес гетероскедастичний, то переходимо до пункту 6, якщо містить тренд, то до кроку 7, інакше до кроку 8.
6	Визначення типу моделі	Визначення типу моделі для опису гетероскедастичності: УАРУГ(p, q), експоненційна УАРУГ(p, q), УАРУГ-М або інша. Після вибору найбільш підходящої моделі переходимо до кроку 9.
7	Визначення способу вилучення або моделювання тренду	Визначення способу вилучення або моделювання тренду включає в себе визначення порядку інтегрованості процесу або можливістю описання тренду поліноміальною функцією, експоненціальною, та іншими функціями. Після вилучення тренду переходимо до кроку 9.
8	Побудова моделі корекції похибок	Для коінтегрованих процесів будується модель корекції похибок і переходимо до кроку 10.
9	Побудова моделі часового ряду	Будується модель часового ряду та обчислюються критерії адекватності отриманої моделі і переходимо до кроку 10.
10	Побудова функції прогнозування	На основі обраної моделі, будується функція прогнозування. Обчислюється прогноз поведінки ряду та визначаються оцінки точності прогнозу.

Для оцінювання якості прогнозів, обчислених за побудованими моделями, запропоновано процедуру автоматизованого вибору кращої прогнозної моделі, у якій для вибору було використано інтегральний критерій якості, який включає 2-5 окремих статистичних критеріїв якості. Процедура автоматизації забезпечує можливість побудови, аналізу та вибору кращої із множини можливих моделей.

4.3 Методика оцінювання адекватності моделей та оцінок прогнозів

У сучасних умовах високої невизначеності та швидкоплинних змін у соціально-економічному середовищі підвищуються вимоги до точності, надійності та адаптивності прогнозних моделей. Прогнозування не може розглядатися лише як побудова функціональної залежності між змінними — воно має супроводжуватись обов'язковим аналізом якості моделі, її здатності відображати реальну динаміку процесів і формувати достовірні прогнози. Цей розділ присвячено розробці інформаційних технологій, які забезпечують комплексне оцінювання якості побудованих моделей і прогнозів. У ньому розглянуто підходи до вибору критеріїв адекватності, реалізації статистичних перевірок, впровадження адаптивних методів оцінювання, а також механізми паралельного контролю за станом і динамікою прогнозованого процесу. Запропоновані рішення є основою для створення гнучких та інтелектуалізованих систем підтримки прийняття рішень в умовах нестабільного середовища.

У складних соціально-економічних системах, що характеризуються високою варіативністю, неоднорідністю та невизначеністю, вибір критеріїв оцінювання є критично важливим етапом для забезпечення достовірності прогнозних моделей. Неправильний або необґрунтований вибір метрики може призвести до некоректної оцінки ефективності моделі, що у свою чергу викликає хибні висновки та управлінські рішення.

Оцінювання якості прогнозних моделей передбачає використання статистичних та інформаційних критеріїв. До основних належать:

Коефіцієнт детермінації (R^2) — показує, яку частку варіації залежної

змінної пояснює модель:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

де y_i — реальні значення, $\hat{y}_i$ — прогнозовані, $\bar{y}$ — середнє значення спостережень.

Середньоквадратична похибка (RMSE) — особливо чутлива до великих відхилень:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Середня абсолютна похибка (MAE) — обчислює середнє абсолютне відхилення:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

AIC (Akaike Information Criterion) та BIC (Bayesian Information Criterion) — використовуються для оцінки відносної якості моделей з урахуванням їх складності:

$$AIC = 2k - 2 \ln(L), \quad BIC = k \ln(n) - 2 \ln(L)$$

де k — кількість параметрів у моделі, L — максимальне значення функції правдоподібності, n — обсяг вибірки.

Умови застосування критеріїв оцінювання залежать від характеру задачі та властивостей даних. Коефіцієнт детермінації (R^2) та середньоквадратична похибка (RMSE) зазвичай використовуються у випадках моделювання безперервних змінних, особливо коли залишки моделі мають нормальний розподіл. Середню абсолютну похибку (MAE) доцільно застосовувати у ситуаціях, коли в даних наявні викиди або необхідно забезпечити однакову вагу

всім відхиленням без посилення впливу великих похибок. Інформаційні критерії, такі як AIC і BIC, є незамінними інструментами при порівнянні моделей — зокрема, в умовах обмеженого обсягу вибірки або коли моделі мають різну кількість пояснювальних змінних. Вибір конкретного критерію повинен враховувати контекст застосування та ціль моделі.

У прикладному сценарії моделювання рівня безробіття в регіонах України було побудовано дві різні моделі: лінійну регресію з чотирма змінними та поліноміальну регресію з десятима змінними. Обидві моделі продемонстрували приблизно однакове значення коефіцієнта детермінації ($R^2 \approx 0.82$), що вказує на подібну здатність пояснювати варіацію залежної змінної. Однак модель з більшою складністю — поліноміальна — мала нижчий показник середньоквадратичної похибки (RMSE), тобто забезпечувала трохи кращу точність прогнозу. Водночас байєсівський інформаційний критерій (BIC) виявився нижчим у простій лінійній моделі, що свідчить про її перевагу з точки зору балансу між якістю апроксимації та складністю. З огляду на це було обґрунтовано відмову від складнішої моделі попри її нижчий RMSE, оскільки вищий BIC вказував на потенційний ризик переобучення та гіршу узагальнюваність на нових даних.

Деталі реалізації оцінювання якості моделей у програмних середовищах надають широкі можливості для автоматизованого аналізу точності прогнозів. У мові програмування Python найбільш поширеним інструментом є бібліотека `scikit-learn`, яка забезпечує обчислення основних метрик за допомогою простих функцій. Зокрема, `r2_score(y_true, y_pred)` використовується для визначення коефіцієнта детермінації (R^2), `mean_squared_error(y_true, y_pred, squared=False)` — для обчислення RMSE (середньоквадратичної похибки, з параметром `squared=False` для квадратного кореня), а `mean_absolute_error(y_true, y_pred)` — для MAE (середньої абсолютної похибки).

Для більш складного аналізу, включно з оцінкою інформаційних критеріїв AIC і BIC, ефективним є використання бібліотеки `statsmodels`, яка дозволяє будувати статистичні моделі регресії з повним доступом до їхніх параметрів,

стандартних похибок, а також обчислення AIC/BIC через атрибути `model.aic` і `model.bic`. Наприклад, після побудови моделі через `OLS()` можна отримати значення інформаційних критеріїв так:

```
import statsmodels.api as sm
model = sm.OLS(y, X).fit()
aic_value = model.aic
bic_value = model.bic
```

У середовищі R аналогічні функції надають пакети `caret`, `lmtest`, `AIC()`, `BIC()`, `summary()` та інші. Наприклад, після побудови моделі за допомогою `lm()` отримати критерії можна так:

```
model <- lm(y ~ x1 + x2, data = dataset)
AIC(model)
BIC(model)
```

Ці інструменти дозволяють швидко та надійно оцінювати якість моделей на основі реальних даних, забезпечуючи високу *reproducibility* досліджень і адаптивність до різних задач соціально-економічного аналізу.

Вибір критеріїв оцінювання є основою коректного аналізу та порівняння прогнозних моделей. Урахування особливостей даних, наявності викидів, типу змінної та цільової функції дозволяє приймати обґрунтовані рішення при виборі найбільш придатної моделі. У соціально-економічних задачах точність, стабільність і простота інтерпретації моделі часто є більш важливими, ніж її математична витонченість.

У задачах прогнозування регіонального розвитку, значення змінних мають складну природу, високий рівень залежностей і вплив багатьох латентних факторів, перевірка статистичних гіпотез виступає ключовим інструментом валідації моделі. Вона дозволяє виявити відхилення від припущень, які лежать в основі лінійного або іншого типу моделювання, і попередити спотворення прогнозу. Перевірка гіпотез охоплює наступні аспекти:

- Гомоскедастичність (постійна дисперсія залишків) — це припущення, що дисперсія залишків моделі є сталою по всьому діапазону значень незалежних

змінних. Порушення цього припущення (тобто наявність гетероскедастичності) означає, що похибки мають різну дисперсію в залежності від рівня предикторів, що може призводити до неефективності оцінок і хибних висновків у статистичному аналізі.

- Тест Бройша–Пагана (Breusch-Pagan) — Цей тест перевіряє, чи є дисперсія залишків функцією від незалежних змінних. Він базується на допоміжній регресії квадратів залишків на ті самі предиктори. Якщо значення p-value низьке (наприклад, < 0.05), це свідчить про наявність гетероскедастичності, тобто порушення гомоскедастичності. Перевагою тесту є його простота та орієнтація на лінійну залежність дисперсії від предикторів.
- Тест Уайта (White) Тест Уайта є більш загальним і не передбачає конкретної форми функціональної залежності дисперсії від змінних. Він здатен виявляти як гетероскедастичність, так і помилки специфікації моделі (наприклад, пропущені змінні чи неправильна форма взаємозв'язку). Тест включає квадрати та добутки змінних у допоміжну регресію, тим самим дозволяє врахувати нелінійні ефекти. Використовується для більш загальної перевірки правильності формулювання моделі, але може бути менш потужним на малих вибірках.
- Автокореляція похибок (або серіальна кореляція) виникає, коли залишки моделі регресії не є незалежними, тобто залишки в поточному періоді пов'язані з залишками в попередніх періодах. Це особливо характерно для часових рядів, де спостереження мають природний порядок у часі.

Порушення припущення про незалежність похибок призводить до: заниження стандартних похибок; переоцінки рівня значущості параметрів; зниження довіри до прогнозів, особливо на тривалих горизонтах.

Тест Дарбіна–Уотсона: Тест Дарбіна–Уотсона є класичним способом виявлення автокореляції першого порядку серед залишків регресійної моделі.

Формула тесту:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

де e_t — залишок у момент часу t , n — кількість спостережень. Інтерпретація результатів: $DW \approx 2$ — відсутність автокореляції. $DW < 2$ — позитивна автокореляція (поширене в економічних даних). $DW > 2$ — негативна автокореляція. Граничні значення залежать від кількості змінних у моделі та обсягу вибірки (табличні критичні значення використовуються для точного рішення).

Приклад: при моделюванні динаміки індексу споживчих цін тест Дарбіна–Уотсона дав значення $DW = 1.35$, що вказує на наявність позитивної автокореляції залишків — тобто похибки в одному місяці систематично схожі з похибками попереднього. У такому разі слід розглянути використання авторегресійних моделей (ARIMA) або додавання лагів у регресію.

Мультиколінеарність — надмірна кореляція між незалежними змінними: Індекс інфляції дисперсії (VIF):

$$VIF_j = \frac{1}{1 - R_j^2}$$

де R_j^2 — коефіцієнт детермінації при регресії x_j на інші змінні.

У моделюванні соціально-економічних процесів важливо визначити, які саме змінні мають статистично значущий вплив на залежну змінну. Для цього проводять t -тести, які дозволяють перевірити гіпотезу про нульовий ефект окремого регресора.

Формулювання гіпотез

Для кожного коефіцієнта β_j регресійної моделі перевіряється нульова гіпотеза:

$$H_0 : \beta_j = 0 \text{ — змінна не має впливу.}$$

$$H_1 : \beta_j \neq 0 \text{ — змінна має статистично значущий вплив.}$$

t -критерій для окремого коефіцієнта:

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

де: $\hat{\beta}_j$ — оцінка регресійного коефіцієнта,

$SE(\hat{\beta}_j)$ — стандартна похибка цієї оцінки.

Отримане значення t порівнюється з критичним значенням із таблиці розподілу Стюдента (t-розподілу) при заданому рівні значущості (зазвичай 0.05).

Інтерпретація результату

Якщо $|t| > t_{кр.}$ або $p\text{-value} < 0.05$ — нульова гіпотеза відкидається, змінна є значущою.

Якщо $|t| \leq t_{кр.}$ або $p\text{-value} > 0.05$ — змінна не має істотного впливу.

Приклад

У моделі прогнозування рівня бідності змінна “освітній рівень” має:

$$\hat{\beta}_j = -1.35,$$

$$SE(\hat{\beta}_j) = 0.42,$$

$$\rightarrow t = -1.35/0.42 \approx -3.21, \text{ що є статистично значущим } (p < 0.01).$$

Таким чином, чим вищий рівень освіти, тим нижчий рівень бідності — і це підтверджено на статистичному рівні.

Показник використовується у перевірці важливості кожного регресора перед включенням у модель. Дає змогу спростити модель, виключивши незначущі змінні. Часто доповнюється побудовою довірчих інтервалів для коефіцієнтів.

Python (statsmodels):

```
import statsmodels.api as sm
```

```
model = sm.OLS(y, X).fit()
```

```
print(model.summary()) # t, p-values, std err
```

R:

```
model <- lm(y ~ x1 + x2, data = dataset)
```

```
summary(model) # t value, Pr(>|t|)
```

Т-критерій — це ключовий інструмент для інтерпретації регресійної моделі, що дозволяє відокремити суттєві предиктори від шуму. Його використання гарантує, що у фінальну модель включено лише ті змінні, які справді впливають на прогнозований показник, що особливо важливо в умовах складних, багатофакторних соціально-економічних систем.

Тест Бройша–Пагана застосовується, якщо на графіку залишків видно змiну дисперсії залежно від рівня предиктора.

Тест Дарбіна–Уотсона — незамінний у моделюванні економічних і соціальних часових рядів.

$VIF > 5$ або 10 вказує на серйозну мультиколінеарність — слід вилучати або трансформувати змінні.

t-тест дозволяє перевірити, чи значущо кожна змінна впливає на результат.

Наприклад, сценарій: моделювання рівня зареєстрованого безробіття в регіонах України.

За результатами аналізу залишків виявлено змінну дисперсію — тест Бройша–Пагана підтвердив гетероскедастичність ($p < 0.05$).

Тест Дарбіна–Уотсона вказав на помірну автокореляцію ($DW = 1.3$, при нормі ≈ 2).

Значення VIF для двох змінних перевищили 8 — було вирішено виключити одну змінну та використати регуляризацію (Ridge).

Деякі змінні мали $p > 0.1$ у t-тесті — їх видалено для покращення узагальнюваності моделі.

Аналіз адекватності статистичних гіпотез дозволяє виявити порушення, які можуть призвести до зміщених або нестабільних оцінок параметрів. Коректне застосування статистичних тестів забезпечує достовірність результатів, особливо в умовах складної структури соціально-економічних процесів. Це підвищує стійкість і ефективність прогнозів, що використовуються в управлінні

та прийнятті рішень на різних рівнях.

Метод адаптивного оцінювання якості побудованих моделей досліджуваних процесів запропоновано для використання для прогнозування процесів регіонального розвитку, в умовах змін, коли класичні статичні моделі можуть втрачати свою ефективність. Тому, виникає потреба в адаптивних підходах, які здатні враховувати нові дані, змінювати ваги факторів і підтримувати точність у динамічному середовищі. Регуляризація та адаптивне оновлення стають ключовими елементами такої стратегії.

Адаптивне оцінювання реалізується через два механізми:

Регуляризоване моделювання Регуляризація дозволяє уникнути перенавчання та стабілізувати модель при високій мультиколінеарності.

Ridge-регресія (L2-регуляризація):

$$\min_{\beta} \sum (y_i - X_i\beta)^2 + \lambda \sum \beta_j^2$$

Штраф зменшує значення коефіцієнтів, але не зануляє їх.

LASSO-регресія (L1-регуляризація):

$$\min_{\beta} \sum (y_i - X_i\beta)^2 + \lambda \sum |\beta_j|$$

Сприяє зануленню деяких коефіцієнтів — виконує відбір змінних. Elastic Net:

$$\min_{\beta} \sum (y_i - X_i\beta)^2 + \lambda_1 \sum |\beta_j| + \lambda_2 \sum \beta_j^2$$

Поєднує L1 і L2-регуляризацію — забезпечує баланс між стабільністю та відбором змінних.

Динамічне оновлення моделей включає автоматичну адаптацію параметрів у міру надходження нових даних (наприклад, через використання ковзного вікна або оновлення коефіцієнтів у реальному часі).

Умови застосування

Ridge — при наявності сильно корельованих предикторів, де важлива

стабільність параметрів.

LASSO — для відбору релевантних змінних, що знижує складність моделі.

Elastic Net — в умовах, коли потрібно збалансувати відбір змінних та контроль над мультиколінеарністю.

Адаптація на потоці — у випадках безперервного оновлення даних (наприклад, моніторинг економічних індикаторів щомісяця).

Наприклад, побудова моделі прогнозу регіонального ВВП на основі понад 30 соціально-економічних показників.

Ridge-регресія стабілізує оцінки в умовах мультиколінеарності (індекс людського розвитку, освітній рівень, демографічна структура).

LASSO дозволяє автоматично відкинути слабо інформативні змінні — з 30 змінних залишилось 12.

Elastic Net зменшує варіацію прогнозу в умовах мінливої інфляції та пандемічного періоду.

Модель автоматично оновлюється кожен квартал, використовуючи нові показники з державних реєстрів.

Метод адаптивного оцінювання дозволяє створювати гнучкі, стійкі до змін моделі, які зберігають свою ефективність у динамічному соціально-економічному середовищі. Регуляризація та адаптивне оновлення дають змогу уникнути перенавчання, зменшити шум і зосередитись на ключових факторах, що впливають на прогноз. Це суттєво підвищує надійність моделей у стратегічному управлінні та політичному плануванні.

Оцінка якості прогнозу є критичним етапом у впровадженні моделі в практичні соціально-економічні системи. Немає сенсу в математично витонченій моделі, якщо її передбачення неточні або нестабільні. Саме тому необхідно використовувати формалізовані метрики точності й процедури перевірки, які гарантують її надійність при застосуванні на нових даних.

Оцінювання якості прогнозу охоплює:

1. Статистичні метрики точності

RMSE (Root Mean Square Error):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Найбільш чутлива до великих похибок.

MAE (Mean Absolute Error):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Усереднює абсолютні відхилення, менш чутлива до викидів.

MAPE (Mean Absolute Percentage Error):

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Вимірює похибку у відсотках, але не застосовується при $y_i = 0$.

2. Методи крос-валідації

Hold-out — поділ вибірки на тренувальну і тестову (наприклад, 70/30):

Простий і швидкий, але залежить від способу розбиття.

K-fold cross-validation — вибірка розбивається на k частин (наприклад, $k=5$):

Навчання відбувається на $k-1$ частинах, а тестування — на залишеній, цикл повторюється.

LOOCV (Leave-One-Out Cross Validation) — граничний випадок $k=n$:

Цей метод використовується при дуже малому обсязі даних, але має високу обчислювальну складність.

Прогнозні інтервали дають оцінку ймовірного діапазону для кожного передбачення:

$$PI = \hat{y}_i \pm z \cdot \hat{\sigma}$$

де z — квантиль нормального розподілу (наприклад, 1.96 для 95%), $\hat{\sigma}$ — оцінена стандартна похибка.

Застосовують, якщо RMSE — коли критичні великі помилки (наприклад, при бюджетуванні).

MAE — якщо похибки однаково важливі незалежно від розміру.

MAPE — для міжрегіональних/міжкраїнних порівнянь у відносних величинах.

K-fold — коли даних небагато, але потрібна стабільна оцінка.

Hold-out — ефективний при великому обсязі даних.

LOOCV — коли вибірка дуже мала, а модель проста.

Прогнозні інтервали — при роботі з чутливими політичними чи економічними рішеннями (наприклад, прогноз безробіття для регіону).

Наприклад: прогноз середньомісячної зарплати на наступний рік.

RMSE = 1500 грн, MAE = 1100 грн, MAPE = 7.3%

K-fold (k=5) показав стабільну похибку на тестових фрагментах.

Прогнозний інтервал для грудня:

$$\hat{y}_{Dec} = 18900, \quad PI_{95\%} = 18900 \pm 1.96 \cdot 650 \Rightarrow [17623; 20177]$$

Це дозволяє аналітику не просто дати точкове передбачення, а й оцінити ризики.

Метод оцінювання якості прогнозу забезпечує всебічну перевірку надійності моделі. Застосування кількох метрик і підходів дозволяє оцінити не тільки середню точність, але й стабільність та межі довіри до результатів. Це є особливо актуальним у прогнозуванні соціально-економічних показників, де точність має прямий вплив на прийняття управлінських рішень, розподіл ресурсів і визначення політики.

Метод одночасного обчислення оцінок стану і динаміки прогнозованого процесу дозволяє не лише передбачати майбутні значення показників, а й мати адекватну оцінку поточного стану системи — особливо коли інформація є неповною або зашумленою. Одним із найбільш ефективних підходів до вирішення таких задач є використання методів фільтрації, зокрема фільтра Калмана, та байєсівського оновлення стану.

Фільтр Калмана — це оптимальний рекурсивний алгоритм, який дозволяє оцінити поточний стан динамічної системи, зважаючи на наявні спостереження та попередні оцінки.

Модель стану:

$$x_t = Ax_{t-1} + Bu_t + w_t$$

Модель спостереження:

$$y_t = Hx_t + v_t$$

де: x_t — вектор стану системи в момент часу t , y_t — вектор спостережень,

A — матриця переходу станів,

H — матриця спостереження,

w_t , v_t — шуми стану та вимірювання відповідно.

Оцінка стану:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - H\hat{x}_{t|t-1})$$

Коефіцієнт підсилення (Kalman gain):

$$K_t = P_{t|t-1}H^T(HP_{t|t-1}H^T + R)^{-1}$$

де:

$\hat{x}_{t|t-1}$ — прогноз стану до спостереження,
 $P_{t|t-1}$ — коваріація похибок стану,

R — коваріація шуму вимірювань.

Умови застосування фільтра Калмана — найефективніший при лінійних динамічних системах із гаусовим шумом.

Байєсівські фільтри (наприклад, Particle Filter) — для нелінійних або нестационарних систем.

Комбіновані моделі стану використовуються у багаторівневих економічних системах, наприклад, коли треба врахувати регіональний і національний рівні.

Наприклад: щотижневий моніторинг стану зайнятості населення у великих містах України. Вхідні дані: часткові спостереження (вибіркові опитування, реєстрація в центрах зайнятості).

За допомогою фільтра Калмана оцінюється повний стан системи: рівень прихованого безробіття, тенденція до скорочення або зростання.

Алгоритм коригує прогноз на основі кожного нового спостереження, забезпечуючи безперервне оновлення.

У складних випадках використовують розширений фільтр Калмана (EKF) або Unscented Kalman Filter (UKF).

Методи одночасного оцінювання стану та динаміки є необхідними в умовах обмеженої або фрагментованої інформації. Вони дозволяють в режимі реального часу адаптувати прогноз до змін середовища. Це робить їх незамінними для оперативного прийняття рішень у складних динамічних соціально-економічних системах, зокрема в кризових ситуаціях, бюджетному плануванні чи моніторингу соціальних ризиків.

Таким чином, сформовано системний підхід до оцінювання якості прогнозного моделювання, що базується на інтеграції класичних статистичних критеріїв, методів перевірки гіпотез, адаптивних механізмів та інтелектуальних технологій. Розроблено інформаційні методи, що дозволяють не лише виявляти слабкі місця моделей, але й оперативно адаптувати їх до змін зовнішнього середовища. Зокрема, продемонстровано ефективність поєднання процедур діагностики регресійних моделей, сценарного аналізу, регуляризації, а також одночасного оцінювання стану і динаміки систем. Запропоновані методики забезпечують високу точність і стабільність прогнозів у складних соціально-економічних умовах, що є передумовою для підвищення ефективності управлінських рішень.

Висновки до розділу 4

1. Запропоновано нові та удосконалено існуючі методи прогнозування, застосовувані у системах підтримки прийняття рішень в управлінні

регіональним розвитком з урахуванням наступних факторів: реформа децентралізації змінила концептуальні підходи до управління регіональним розвитком – територіальні громади отримали нові можливості та повноваження, а значна їх кількість об'єднались, утворивши нові громади – кардинально змінилась структура регіональної системи, а методичні підходи до розроблення стратегій, планів, програм регіонального розвитку, основою яких є виключно ретроспективні дані, за таких умов не забезпечують високу якість прогнозів.

2. Встановлено, що потребує удосконалення як організаційна складова аналітичної роботи, так і методичні підходи до побудови прогнозів.
3. Запропоновано підхід до побудови ансамблів математичних моделей регіонального розвитку, який на відміну від наявних має модифіковану комплексну структуру, збільшує адекватність та точність прогнозування сценаріїв зміни соціально-еколого-економічних систем, що дає змогу формувати надійні прогнози.
4. Удосконалено метод прогнозування на основі адаптивного підходу до моделювання та прогнозування з використанням регресійних, узагальнених лінійних і ймовірнісно-статистичних моделей у формі мереж Байєса, який, на відміну від наявних, враховує можливі структурно-параметричні невизначеності моделей і забезпечує коректний формальний опис причинно-наслідкових зв'язків, що дозволить підвищити точність прогнозів.
5. Враховуючи особливості досліджуваних процесів, запропоновано критерії адекватності моделей та оцінки якості прогнозів, які, на відміну від відомих, забезпечать комплексну перевірку проміжних і кінцевих результатів моделювання, та підвищить достовірність прогнозів і якість підтримки прийняття рішень у процесах регіонального розвитку

РОЗДІЛ 5

МЕТОДОЛОГІЯ ПОБУДОВИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ У СИСТЕМІ УПРАВЛІННЯ РЕГІОНАЛЬНИМ РОЗВИТКОМ

5.1 Архітектурні рішення щодо створення інформаційних технологій

Для вирішення задач дослідження запропоновано методологію побудови інформаційних технологій для підтримки прийняття рішень у системі управління регіональним розвитком, яка побудована виходячи з особливостей наявної організації структури управління регіонами та громадами, просторового розташування складових системи. Розроблювані інформаційні технології поєднано у інформаційно-керуючій системі, побудованій за мережецентричним принципом, взаємодія об'єктів та процесів управління якої відбувається у єдиному інформаційному просторі (рис. 5.1).

Мережецентричне середовище (єдиний інформаційний простір)

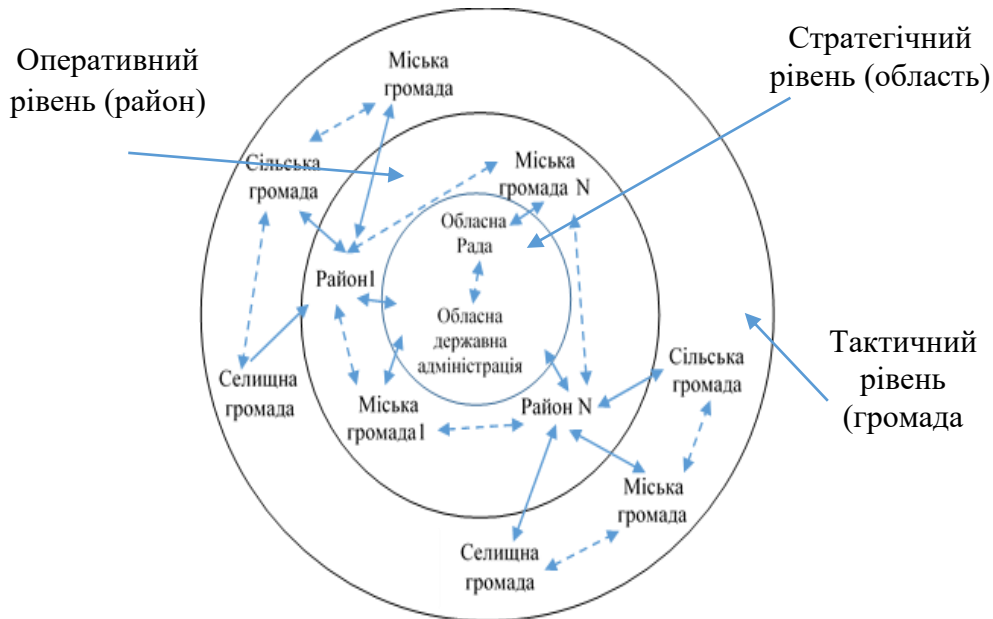


Рисунок 5.1 – Схема взаємодії об'єктів у системі управління регіональним
розвитком

Використання єдиного інформаційного простору, зокрема, загальних бази даних та базу знань, у системі забезпечується ситуаційна обізнаність для всіх рівнів регіонального управління.

Архітектура системи є багаторівневою із вертикальними та горизонтальними зв'язками – являє собою масштабований адаптивний мережевий гіперграф великої розмірності, що дозволяє зв'язати різномірні компоненти існуючої системи практично «безшовно».

Інформаційні технології для підтримки прийняття рішень управління регіональним розвитком пропонується використовувати у складі єдиної інформаційно-аналітичної платформи управління регіональним розвитком, до складу якої пропонується олучити Національну електронну систему управління відбудовою DREAM, геоінформаційну систему регіонального розвитку. Документообіг має забезпечувати система електронної взаємодії органів виконавчої влади. Має також забезпечуватись електронну взаємодію між системою та державними інформаційними ресурсами, системою та електронними реєстрами, системою та інформаційними іншими державними інформаційними системами засобами системи «Трембіта».

Формування єдиного інформаційного простору, ставить вимоги до інтероперабельності складових системи та забезпечення оброблення значних обсягів даних, їх зберігання та спільного використання на різних рівнях органів державного управління та місцевого самоврядування, надавати доступ до відкритих даних широкому загалу. Тому, для якісної реалізації поставлених задач, пропонується використовувати апаратно-програмну платформу Oracle [49]. Дане програмне забезпечення добре інтегрується із програмними продуктами інших вендорів, середовищами розробки програмного забезпечення, і за своїми характеристика є оптимальним для розв'язання більшості задач у системі підтримки прийняття рішень в управлінні регіональним розвитком. Кращим варіантом реалізації аналітичної складової, зокрема, економетричного, математичного моделювання, інтелектуального аналізу даних, використання штучного інтелекту, тощо та представлення результатів у формі документів та

дашбордів, пропонується поєднання використання спеціалізованого програмного забезпечення Oracle BI та аналітичної платформи компанії SAS Institute [50]. Таке поєднання забезпечить надійну програмно-апаратну платформу та потужні аналітичні засоби, до яких легко інтегруються більшість сучасних середовищ розробки програмного забезпечення. Що додатково створить переваги доопрацювання програмного забезпечення системи до потреб користувачів чи задач, які розв'язуються.

Схема системи підтримки прийняття рішень на базі розроблених інформаційних технологій прогнозування регіонального розвитку представлена на рис. 5.

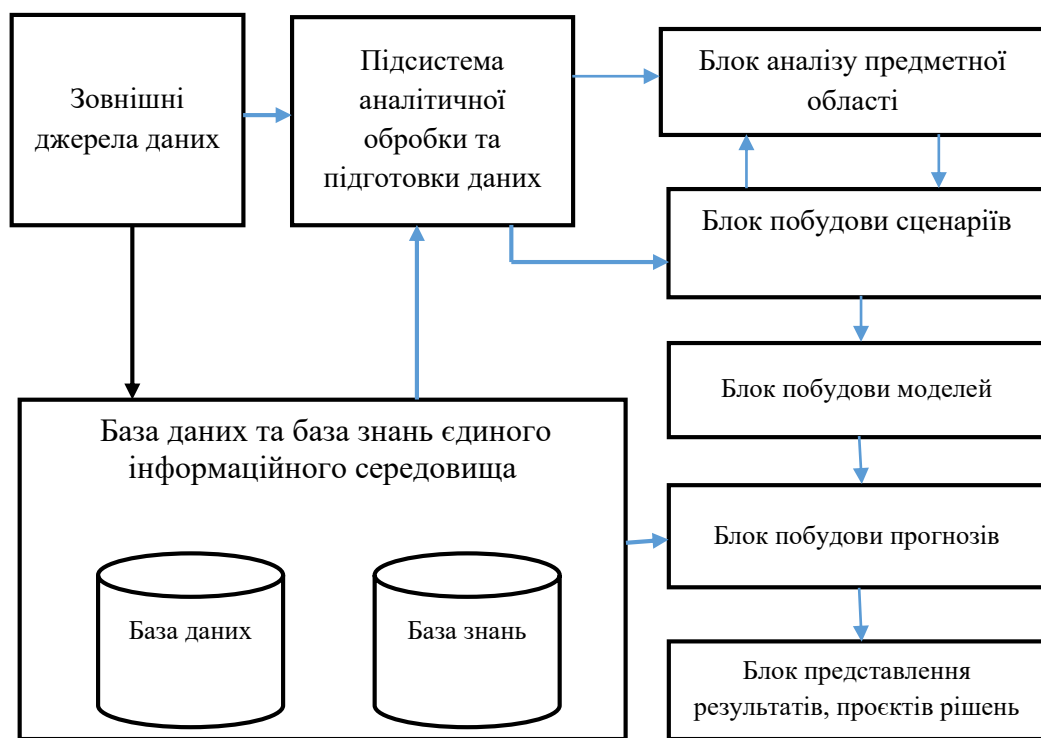


Рисунок 5.2 – Загальна схема застосування інформаційних технологій у системі підтримки прийняття рішень в управлінні регіональним розвитком

Особливістю системи є те, що у підсистемі аналітичної обробки та підготовки даних передбачено виконання SWOT та PEST – аналізу, а побудова моделей може виконуватись як в рамках конкретного сценарію (сценаріїв), так і до сценарію можна долучити необхідні моделі, в тому числі економетричні,

математичні, геопросторові, когнітивні та ймовірісно-статистичні, тощо. База знань системи використовується як при побудові сценаріїв, та і в процесі формування проекту рішення. База моделей нарощуватиметься у процесі використання. Для розв'язання задачі побудови моделей та прогнозів, засоби бізнес-аналітики, компанії Oracle пропонується доповнити програмним забезпеченням компанії SAS Institute, яке реалізує широкий набір аналітичних інструментів, має зручний інтерфейс, який дозволяє користувачу власноруч будувати потрібну технологічну схему (без написання програмного коду) та завантажувати необхідні дані. Для рівня громади і регіону пропонується використовувати типовий набір моделей за розробленими типовими схемами аналітичних процесів, які за потреби, можуть бути адаптовані до особливостей конкретних вирішуваних задач.

Дані для побудови моделей та прогнозів за потреби, можуть бути завантажені як структуровані (таблиці, файли баз даних, текстові файли, тощо), так і не структуровані (текст звернень, скановані зображення, фото, тощо). Використання засобів Oracle BI дозволяє інтегрувати та обробляти структуровані та неструктуровані дані для безпосереднього використання при побудові моделей та прогнозів та обробляти та накопичувати у базі даних або хмарних сховищах.. Завантаження даних в систему побудови моделей та прогнозів може здійснюється безпосередньо з сховищ системи - бази даних та хмарного сховища. Або, дані можуть бути завантажені, використовуючи API зовнішніх джерел. Наприклад, інформація з бази даних бюджетів територіальних громад, відомості з державних реєстрів, дані державних установ, міністерств, відомств – використовуючи систему електронної взаємодії «Трембіта» і підготовлені до використання у системі моделювання та прогнозування. Саме цей етап роботи системи є чи не найскладнішим з точки зору забезпечення її інтеоперабельності. На цьому етапі на семантичному рівні моделі інтеоперабельності має бути забезпечена узгодженість змістовного наповнення інформації стосовно функціонування громад та регіонів з наявними статистичними даними та даними звітності територіальних громад, підприємств, установ таким чином, що

уникнути дублювання, надлишковості даних з одного боку, а з іншого – отримати відомості, що в повній мірі описують стан та динаміку регіонального розвитку. На організаційному рівні має бути узгоджена та закріплена нормативними актами, угодами, тощо, взаємодія між різними гілками публічного та державного управління, а також взаємодія бізнесу та держави, організована робота із зверненнями громадян та розроблена методика їх аналізу та його використання у роботі органів місцевого самоврядування та публічного управління. На технічному рівні пропонується використання єдиної апаратно-програмної платформи та узгодження протоколів, процедур та форматів представлення та передавання даних, призначених до використання у системі управління регіональним розвитком.

Оброблення даних про стан та динаміку територіальних громад виконується на різних рівнях ієрархії «громада-регіон-держава». На кожному з рівнів обробляється та частина інформації, яка відповідає рівню управлінських рішень. На рівні громади, звичайно, найчастіше використовуваними є дані щодо її функціонування, однак, для виявлення «вузьких місць», визначення перспектив розвитку, на рівні громади доцільно порівнювати показники соціально-економічного розвитку громади з кращими для громад, аналогічного типу (для типологізації громад та регіонів існують і законодавчо визначені методики, але й можуть бути розроблені власні). Для цього необхідно забезпечити доступ до результатів узагальнення даних, які може отримати фахівець відповідного рівня управління, звернувшись з запитом до бази даних. У системі передбачений модуль роботи з даними інших громад та регіонів, скориставшись яким, користувач може власноруч задати параметри для класифікації об'єктів або вибрати наявний варіант типологізації.

Оскільки, значна увага приділяється роботі з даними, то пропонується для розв'язання цієї задачі застосувати конвергентний підхід до організації збору, оброблення та зберігання даних. Тобто, має бути формується система, центральною частиною якої є центр оброблення даних, а складовими - система обміну даними між компонентами, хмарні сервіси. Дані для аналізу

використовуватимуться не лише з аналітичного сховища (пропонується інтеграція Nadoor та ETL/ELT, хмарні сховища), а й в результаті виконання API-запитів, в тому числі, коли потрібно здійснювати обробку даних у режимі реального часу. Для представлення результатів оброблення даних у формі документів та дашбордів, пропонується використання спеціалізованого програмного забезпечення.

В системі реалізовано потокову обробку даних, що забезпечує безперервність збору та аналізу інформації - дозволяє оперативно реагувати на події в реальному часі. Основними складовими пропонованої технології потокової обробки є брокер повідомлень, який організовує надійну доставку даних у режимі реального часу від сенсорів, реєстрів державних систем (DREAM, «Трембіта»), а також з відкритих джерел в Інтернет, що створює єдиний транспортний шар для подій у системі «група – район – область». Неперервна обробка даних: агрегація, трансформація та очищення потоків даних, дає можливість вже на попередніх етапах оброблення інформації використовувати алгоритми машинного навчання для виявлення трендів або аномалій. Інтеграція стрімінгової аналітики у системи підтримки прийняття рішень забезпечує зниження латентності процесів управління, оскільки результати аналізу стають доступними практично одразу після надходження даних. Це створює можливість формування актуальних дашбордів для управлінців різних рівнів, що підвищує прозорість та оперативність управлінських дій.

Наприклад, для використання у аналітичній роботі даних з інтегрованою інформаційно-аналітичною системою «Прозорий бюджет» запропоновано підсистему аналізу місцевих бюджетів, яка дозволяє не тільки оцінювати спроможність громади, а й порівняти з іншими громадами такого ж типу або дослідити досвід кращих громад. Для виконання аналітичних та прогнозних функцій модуля використано байєсівський підхід до аналізу даних, умовні багатовимірні розподіли, регресійні моделі, моделі стохастичної волатильності.

Система включає в себе клієнтську частину, доступну для користувачів, яка дозволяє визначати набір даних та встановлювати кінцеві цілі через інтерфейс користувача (UI), реалізований за допомогою бібліотеки ReactJS та мови програмування Python. Взаємодія між клієнтською та серверною частиною системи здійснюється через API за допомогою HTTP-запитів.

Користувачі взаємодіють з клієнтською частиною, використовуючи веб-інтерфейс, де вони можуть вибирати параметри, налаштування та ініціювати запити. Ці запити передаються серверній частині системи через API. Серверна частина обробляє ці запити, взаємодіє з базою даних PostgreSQL для отримання необхідної інформації, проводить аналіз даних та генерує відповіді для клієнтів.

API визначає структуру та формат запитів і відповідей, що дозволяє коректно обмінюватися інформацією між клієнтом і сервером. Цей механізм взаємодії через HTTP-запити дозволяє користувачам отримувати результати аналізу даних та зберігати інтермедіатні значення в системі. У результаті цієї взаємодії користувачі можуть виконувати різноманітні завдання та отримувати необхідну інформацію з системи через API-запити.

Крім того, завдяки поєднанню сценарного аналізу з методами прогнозування на основі LSTM, байєсівських моделей чи когнітивних підходів, система здатна швидко реагувати на кризові ситуації та адаптуватися до умов невизначеності.

Наприклад, на етапі збору даних інформація надходить від сенсорів екологічного моніторингу (якість повітря, рівень забруднення води, шумове навантаження), статистичних реєстрів (зайнятість населення, рівень доходів, демографічні показники) та відкритих онлайн-джерел (соціальні мережі, відгуки громадян, новини). Усі ці дані в режимі реального часу передаються у брокер повідомлень, що забезпечує їх централізовану та надійну доставку. Подальша потокова обробка здійснюється у модулі стрімінгової обробки, де дані проходять очищення, агрегуються на рівні громад та районів і аналізуються з використанням часових моделей. Алгоритми глибинного навчання, зокрема LSTM, формують прогнози щодо можливих змін соціально-економічних

показників (наприклад, зростання рівня безробіття чи міграційних потоків) та соціально-екологічних ризиків (наприклад, погіршення стану довкілля у зв'язку з інтенсивним промисловим навантаженням).

На етапі аналітичної обробки, результати інтегруються у аналітичному модулі системи, де виконується сценарний аналіз, пропонуються варіанти рішень, що ефективно у процесі реагування на кризові ситуації (табл. 5.1).

Таблиця 5.1 - Схема формування рекомендацій у процесі прийняття рішень у кризовій ситуації

Етап	Зміст / Методи
Початок	Старт процесу прийняття рішень
Вхідні дані	Структуровані: статистика, державні реєстри Неструктуровані: звітність, тексти, веб, соціальні мережі Стрімінгові сервіси: Kafka, Spark Streaming, Flink
База знань та даних	Центр обробки даних: Hadoop, хмарні сервіси (AWS, GCP, Azure), API Інтеграція реєстрів: DREAM, Трембіта Аналітичні сховища: ETL/ELT, Data Lake
Попередня обробка	Заповнення пропусків: KNN imputation, ЕМ-алгоритми Очищення, нормалізація - Часові ряди: ARIMA, Prophet, LSTM - Текстова аналітика: TF-IDF, Word2Vec, BERT, LDA - Кластеризація: k-means, DBSCAN, ієрархічна - Когнітивне моделювання: когнітивні карти
Генерація кандидатів-сценаріїв	Ансамблі моделей: регресійні, байєсівські мережі, імовірнісні моделі (HMM, Monte Carlo), LSTM/GRU, GNN - Формування альтернатив розвитку (оптимістичний, базовий, песимістичний сценарії)
Оцінка та ранжування	Методи багатокритеріального аналізу: AHP, TOPSIS, ELECTRE, PROMETHEE - Критерії: економічні, соціальні, екологічні, інфраструктурні - Оптимізація: генетичні алгоритми, swarm intelligence (PSO), simulated annealing - Функції корисності
Перевірка відповідності вимогам	Адаптивність - Стійкість до ризиків (stress-testing, what-if) - Узгодженість зі стратегіями розвитку
Переоцінка (зміна рейтингу)	Моделювання ризиків: Bayesian risk models, нечіткі множини, VaR - Адаптація: reinforcement learning (Q-learning, DQN) - Автоматизоване реагування (II)
Формування готових рекомендацій	Варіанти сценаріїв для громад, регіонів, держави. Візуалізація: Power BI, Tableau, Grafana. Документи, аналітичні звіти. Зворотний зв'язок: опитування, чат-боти, електронні сервіси
Кінець	Вибір оптимального рішення з множини варіантів (з урахуванням даних та експертних оцінок)

Перевагою такого підходу те, що він може бути використаний як для стратегічних прогнозів, так і для швидкого прийняття рішень у випадку загрози криз. Результати подаються у вигляді дашбордів, що відображають прогнозовані показники розвитку соціально-еколого-економічних процесів, а також потенційні ризики для громади чи регіону.

5.2 Розроблення прототипу системи підтримки прийняття рішень, що використовує розроблені інформаційні технології

Розроблені інформаційні технології поєднані у систему підтримки прийняття рішень управління регіональним розвитком. Метод синтезу інформаційних технологій представлений на рис. 5.3.

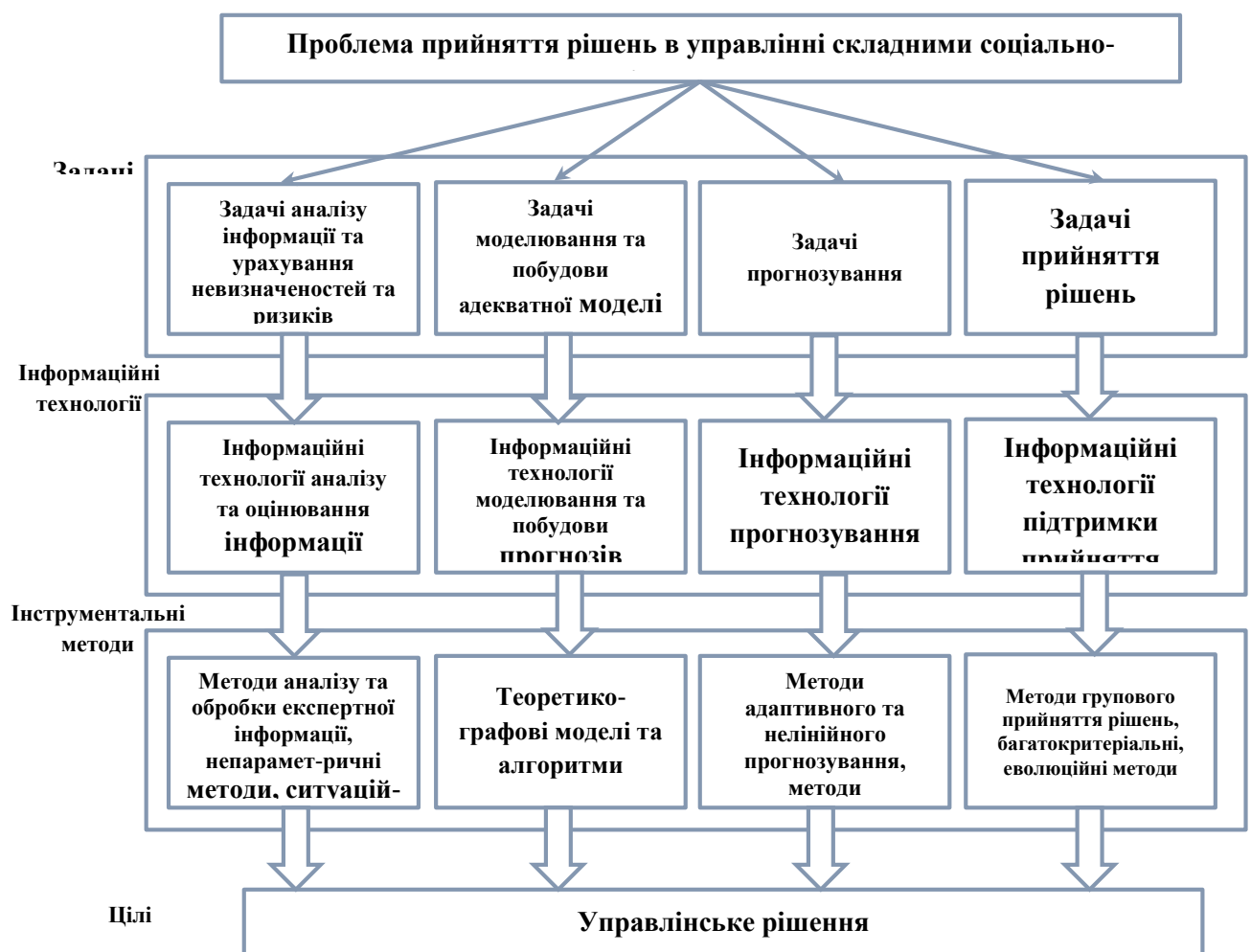


Рисунок 5.3 – Метод синтезу інформаційних технологій

Центральними компонентами розробленої системи є база даних та база знань, розпорядником яких має бути держава. Основна частина (ядро системи) належить державі, містить програмні модулі, реєстри, засоби захисту інформації, які забезпечують накопичення, зберігання, оброблення інформації, отримуваної як всередині системи, так і з зовнішніх джерел, зокрема, державними інформаційними ресурсами, іншими інформаційними системами. Інформаційні системи територіальних громад – є ініфікованими, належать державі та забезпечують доступ до бази даних та знань системи. Інші електронні ресурси – це програмні продукти користувачів (це бізнес, громадські організації, дорадницькі структури, агенції регіонального розвитку) спеціалізоване програмне забезпечення для організації взаємодії з основною («державною») частиною системи. Взаємодія таких додатків з системою має здійснюватись на договорних засадах. Система реалізована за мережецентричним підходом, що передбачає створення єдиного інформаційного простору, що є перевагою для всіх користувачів системи, оскільки найважливіші дані зберігаються у центральній базі даних, користувачі можуть завантажувати власні дані, користуючись відповідним інтерфейсом. Система захисту інформації має відповідати вимогам чинного законодавства України, вітчизняним та міжнародним стандартам.

Схема сегмента системи наведена на рис. 5.4

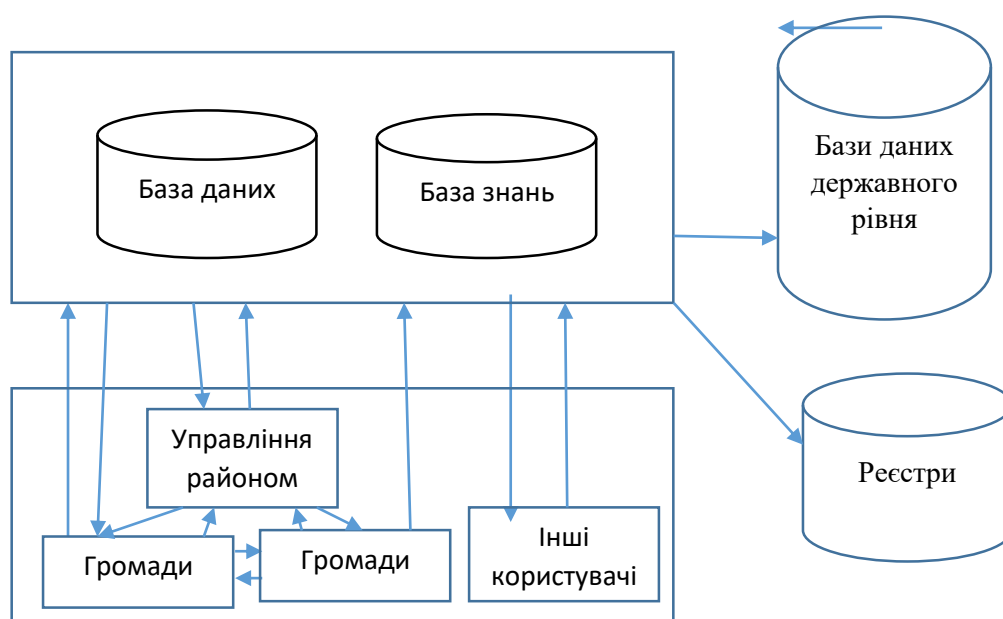


Рисунок 5.4 – Схема взаємодії об'єктів системи (район-область)

Роботу системи можна представити у декілька етапів.

Етап 1. Формування вихідних умов

Вікно «Ініціалізація налаштувань»—це перший крок для налаштування інформаційної технології підтримки прийняття рішень. Воно збирає мінімально достатній набір параметрів, щоб сформувати робочий контекст: геопросторові та часові межі аналізу, цільові орієнтири, релевантну нормативну базу, а також початкові джерела даних (локальні файли або зовнішні API-ресурси). Декомпозиція на шість логічних компонентів забезпечує швидкий старт без перевантаження користувача деталями та водночас гарантує цілісність подальших етапів обробки даних.

На рис. 5.5 представлено узагальнену структуру вікна «Ініціалізація налаштувань», яка окреслює ключові UI-елементи першого кроку інтерфейсу.

Рисунок 5.5 – Інтерфейс вікна

У таблиці 5.1 систематизовано компоненти за типом керуючого елемента та функціональним призначенням.

Таблиця 5.2 – Опис компонентів вікна

Компонент	Тип	Призначення
Вихідні умови	<Select> (drop-down)	Вибір базового контексту: країна, область, район, громада, місто.
Ціль	<Textarea> / <Input>	Введення стратегічних або операційних цілей аналізу. Підтримка Markdown.
Нормативна база	<CheckboxGroup>	Вибір нормативних документів (Постанова, Стратегія, Закон) з коротким описом при наведенні.
Файл	<Upload.Dragger>	Завантаження CSV/Excel/JSON файлів із прев'ю перших рядків.
Джерело	<CheckboxGroup + Submenu>	Вибір зовнішніх API-ресурсів (Держстат, Дія, Trembita, DREAM, OpenBudget) з можливістю авторизації.
Рівень деталізації	<CheckboxGroup>	Визначення масштабу аналізу: держава, область, район, громада, місто, мікрорайон.
Тип аналізу	<RadioGroup>	Вибір логіки роботи системи: прогнозування, сценарний аналіз, моніторинг.
Часові рамки	<Slider + Input>	Встановлення періоду аналізу (наприклад, 2000–2025).
Горизонт прогнозування	<RadioGroup>	Вибір горизонту прогнозування: короткостроковий, середньостроковий, довгостроковий.
Кнопка «Продовжити»	<Button type="primary">	Активна, якщо мінімальні умови налаштування виконані (обрано джерело/файл та вихідні умови).

Вікно ініціалізації виконує роль «точки входу», що консолідує ключові атрибути майбутнього аналізу та забезпечує контроль якості вхідних даних ще до запуску важких обчислювальних процесів. Чітко сформульовані умови

активації кнопки «Продовжити» запобігають неповному налаштуванню, а продуманий UI підвищує швидкість роботи та зменшує кількість помилок користувача.

Етап 2. Збір і попередня обробка даних

Процеси «Збору і попередньої обробки даних» охоплюють етап, на якому необроблені дані завантажуються, очищуються та структуруються для подальшого аналізу. Вікно «Обробка даних» слугує інтерактивним центром керування цим процесом: користувач обирає джерело даних, задає типи предобробки, визначає середовище зберігання та, за потреби, вибирає хмарний сервіс. Така конфігурація дозволяє балансувати між гнучкістю ad-hoc підходу й стандартизованими пайплайнами, забезпечуючи відтворюваність і контроль якості даних. На рисунку 5.6 наведено інтерфейс етапу.

The screenshot shows a web interface for data processing configuration, divided into two main panels: 'Обробка даних' (Data Processing) and 'Додаткові налаштування' (Additional Settings).

Обробка даних (Data Processing):

- Джерело даних (Data Source):** Two buttons labeled 'Додайте файл' (Add file).
- Типи обробки (Processing Types):** A list of checkboxes:
 - ☐ Первинна обробка (Ingest)
 - ☐ Очищення даних (Clean)
 - ☐ Збагачення (Enrichment)
 - ☐ Трансформація (Transform)
 - ☐ Інтеграція (Join)
 - ☐ Валідація (Validation)
 - ☐ Анонімізація (Anonymization)
 - ☐ NLP / AI-аналітика
- Зберігання (Storage):**
 - Radio buttons: ☒ Хмара (Cloud), ☐ Локально (Locally), ☐ Гібридно (Hybrid).
 - Dropdown menu: AWS S3 (selected), GCP Storage, Azure Blob.

Додаткові налаштування (Additional Settings):

- Фічерінг (Feature Engineering):**
 - ☐ Ковзні середні (Rolling averages)
 - ☐ Лаги (Lags)
 - ☐ Індикатори тренду (Trend indicators)
 - ☐ Автовідбір ознак (Automatic feature selection)
- NLP-обробка (NLP Processing):**
 - ☐ Токенізація (Tokenization)
 - ☐ Векторизація (TF-IDF, Word2Vec, BERT)
 - ☐ Класифікація текстів (Text classification)
- Data Quality Report:**
 - ☐ Completeness
 - ☐ Validity
 - ☐ Duplicates check
 - ☐ Генерація PDF-звіт (Generate PDF report)
- Автоматизація / AI-допомога (Automation / AI Assistance):**
 - ☐ AutoML preprocessing pipeline (AI підбирає процедури) (AI selects procedures)
 - ☐ Auto Data Profiling (AI робить огляд набору даних і пропонує потрібні трансформації) (AI reviews the dataset and suggests necessary transformations)

A 'Продовжити' (Continue) button is located at the bottom right.

Рисунок 5.6 – Вікно «Збору і попередньої обробки даних»

У таблиці 5.4 наведено перелік основних елементів вікна «Обробка даних». Користувач може налаштувати джерело, типи первинної обробки, спосіб зберігання та запуск подальшого data-pipeline. Кожен компонент таблиці вказує

тип UI-контролю та його функціональне призначення, формуючи єдину точку керування процесом інжесту й очищення даних.

Таблиця 5.3 Основні елементи вікна «Обробка даних»

Компонент	Тип / елемент	Функція
Джерело даних	<Select>	Обрати попередньо зареєстроване джерело (CSV, API, DB-конектор).
Upload-блок	<Upload.Dragger>	Додати локальний файл для завантаження даних.
Процедури обробки	<Checkbox.Group>	- Первинна обробка (Ingest) - Очищення даних (Clean) - Збагачення (Enrichment) - Трансформація (Transform) - Інтеграція (Join) - Валідація (Validation) - Анонімізація (Anonymization) - NLP / AI-аналітика
Зберігання	<Radio.Group>	Вибір способу зберігання: Хмара / Локально / Гібридно.
Вибір сервісу	<Select>	Активується, якщо обрано «Хмара» або «Гібридно». Сервіси: AWS S3, GCP Storage, Azure Blob.
Фічерінг (Feature Engineering)	<Checkbox.Group>	- Ковзні середні - Лаги - Індикатори тренду - Автовідбір ознак
NLP-обробка	<Checkbox.Group>	- Токенізація - Векторизація (TF-IDF, Word2Vec, BERT) - Класифікація текстів
Data Quality Report	<Checkbox.Group>	- Completeness - Validity - Duplicates check - Генерація PDF-звіту
Автоматизація / AI-допомога	<Checkbox.Group>	- AutoML preprocessing pipeline (AI підбирає процедури) - Auto Data Profiling (AI аналізує набір даних і пропонує трансформації)
Кнопка «Продовжити»	<Button type="primary">	Активація переходу до наступного етапу (розклад чи запуск pipeline).

Після заповнення ключових полів і підтвердження вибраних процедур кнопка «Продовжити» ініціює створення або запуск data-pipeline, який автоматизує інжест, очищення та збереження набору у вибраному середовищі. Це завершує первинну підготовку даних, формуючи якісний фундамент для подальших трансформацій, валідації та аналітичних обчислень.

Етап 3. Інтеграційна платформа

У вікні «Інтеграційна платформа» відображено, як система уніфікує керування всією інфраструктурою: від дата-центру і обчислювальних кластерів до мережевих вузлів та API-шлюзів. Користувач одразу бачить топологію на інтерактивному Canvas-графі, швидко перемикається між шарами через лівий сайдбар і редагує властивості вибраних ресурсів у правій панелі. Завдяки drag-and-drop, zoom, multi-select і «гарячим» клавішам інтерфейс поєднує візуальну наочність із продуктивними діями «в один клік».

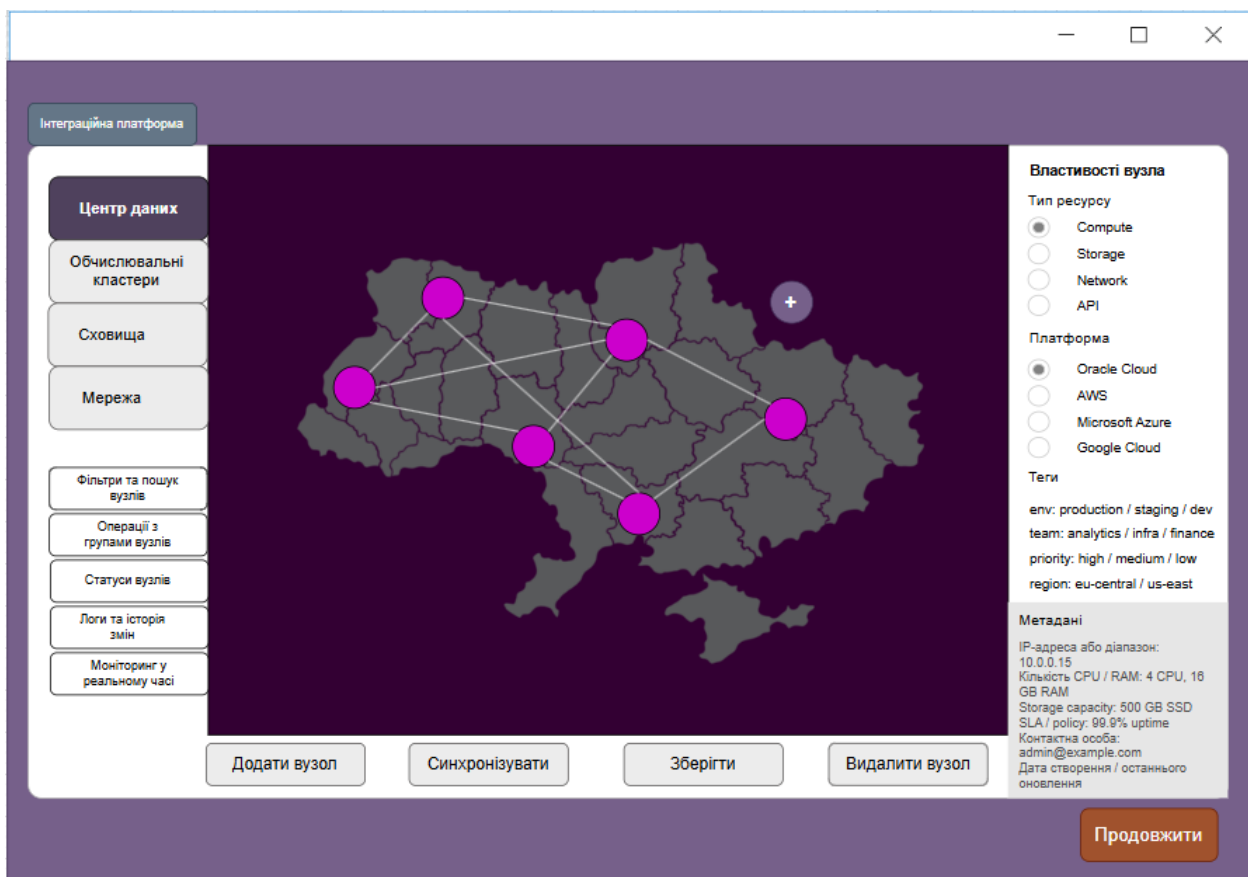


Рисунок 5.7 - Вікно «Інтеграційна платформа»

У таблиця 5.4 наведено структурований перелік ключових елементів вікна «Інтеграційна платформа». Таблиця дає швидкий огляд того, як

організовано навігацію між шарами інфраструктури, відображення топології на Canvas-графі та керування властивостями окремих вузлів. Додаткові UX-деталі — індикатори синхронізації, drag-edge і гарячі клавіші — підсилюють інтерактивність, забезпечуючи оперативне моделювання й актуалізацію конфігурацій у зв'язці з Oracle Cloud Infrastructure.

Таблиця 5.4 – Узагальнений перелік ключових елементів вікна «Інтеграційна платформа»

Компонент	Призначення
Лівий сайдбар	Швидка навігація по шарах інфраструктури: Центр даних, Обчислювальні кластери, Сховища, Мережа. Додаткові опції: Фільтри та пошук вузлів, Операції з групами вузлів, Статуси вузлів, Логи та історія змін, Моніторинг у реальному часі.
Canvas-граф (мапа України)	Візуальне представлення топології дата-центрів у вигляді вузлів на карті з лініями-зв'язками. Підтримує drag-n-drop, zoom, multi-select та додавання нових вузлів через «+».
Права панель «Властивості вузла»	Динамічна форма для редагування вибраного вузла: вибір типу ресурсу (Compute, Storage, Network, API), вибір платформи (Oracle Cloud, AWS, Microsoft Azure, Google Cloud), додавання тегів (env, team, priority, region) та метаданих (IP-адреси, CPU/RAM, storage, SLA, контактна особа, дата створення/оновлення).
Нижня action-bar	Керування діями: Додати вузол → відкриває модальне вікно для конфігурації; Синхронізувати → звіряє локальну топологію з живим середовищем; Зберегти → коміт змін у репозиторій конфігурацій; Видалити вузол → очищає виділений елемент.
Кнопка «Продовжити» (праворуч унизу)	Перехід до наступного етапу інтеграційного процесу після збереження конфігурації.

У таблиці 5.5 узагальнено основні поля правої панелі «Властивості вузла», що використовуються для опису ресурсів у інтеграційній платформі.

Таблиця 5.5 - Опис ресурсів у інтеграційній платформі

Поле	Можливі значення	Опис / призначення
Тип ресурсу	Compute / Storage / Network / API	Визначає, до якого класу належить вузол. • <i>Compute</i> — віртуальні машини, контейнери, bare metal. • <i>Storage</i> — блокове, файлове чи об'єктне сховище. • <i>Network</i> — маршрутизатор, VPN-шлюз, балансувальник. • <i>API</i> — вузол доступу до сервісів (REST, GraphQL, gRPC).
Платформа	Oracle Cloud / AWS / Microsoft Azure / Google Cloud	Хмарна або локальна платформа, на якій розгорнуто вузол. Вибір впливає на доступні параметри метаданих (наприклад, у AWS можна зберегти ID інстансу, у Azure — ресурсну групу).
Теги	env: production / staging / devteam: analytics / infra / financepriority: high / medium / lowregion: eu-central / us-east	Додаткові логічні мітки, що дозволяють групувати й швидко шукати вузли. Використовуються для фільтрації, контролю доступу та автоматизації сценаріїв (наприклад, оновлення тільки для вузлів з тегом env=dev).
Метадані	• IP-адреса або діапазон: 10.0.0.15 • CPU/RAM: 4 CPU, 16 GB RAM • Storage capacity: 500 GB SSD • SLA/Policy: 99.9% uptime • Контактна особа: admin@example.com • Дата створення / останнього оновлення	Технічні або бізнес-атрибути вузла. Використовуються для документування й інтеграції з моніторингом. Дозволяють зберігати як структуровану інформацію (ресурси, SLA), так і довільні примітки.

Вона охоплює класифікацію вузла за типом і платформою, логічні мітки для групування (теги) та технічні атрибути (метадані), які допомагають у документуванні й моніторингу. Така структура забезпечує швидке налаштування та уніфікований підхід до керування інфраструктурними об'єктами.

Набір інструментів нижньої action-bar (додавання вузлів, синхронізація з Oracle Cloud, збереження у репозиторій) разом із індикаторами стану синхронізації та фільтрами шарів забезпечує повний цикл роботи з інтеграційною схемою — від моделювання до контролю актуальності. Таким чином, користувач отримує єдину точку правди, що скорочує час на оновлення конфігурацій і мінімізує ризик розбіжностей між проєктною моделлю та живим середовищем.

Етап 4. Аналітична підсистема

Повноекранне вікно «Аналітична підсистема» (рис. 5.8) поєднує розгалужену навігацію (Home → Breadcrumb → Профіль), тематичний лівий сайдбар (SWOT, PEST, Text/Geo-mining, Моделі, What-if) та гнучку data-панель із вибором факторів і перемикачем

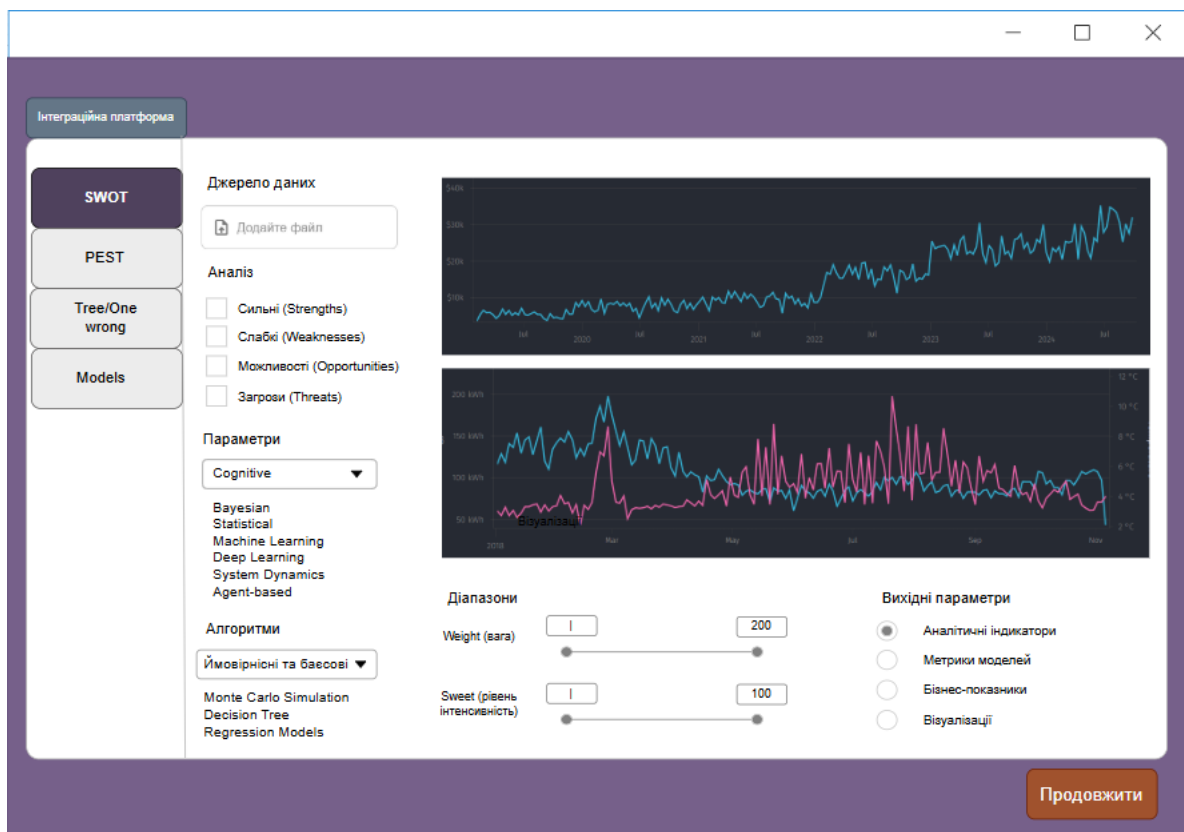


Рисунок 5.8 – Приклад інтерфейсу аналітичної підсистеми»

«Geospatial data» - центральна область відведена під інтерактивну SWOT-матрицю із міні-картою, що активується при ввімкненні геоданих, тоді як секція Output & Insights відображає сценарний графік, аналіз чутливості й рейтинг ключових факторів. Параметри моделі (алгоритм, ваги) та ієрархічний Model repository винесені праворуч, забезпечуючи швидкий доступ до налаштувань і версій.

Повноекранний макет аналітичної підсистеми об'єднує навігацію (Home → Breadcrumb → Профіль), багатофункціональний лівий сайдбар (SWOT, PEST, Text/Geo-mining, Моделі, What-if) та data-панель для вибору факторів і геоданих. У центрі — інтерактивна SWOT-матриця з міні-картою, праворуч — **Output & Insights** (графік, чутливість, топ-фактори) й панель параметрів із репозиторієм моделей, що дозволяє одночасно керувати даними, моделями й результатами (таблиця 5.6).

Таблиця 5.6 - Опис ресурсів у інтеграційній платформі

Параметр	Тип даних	Опис
Джерело даних	File Upload (CSV/XLS/JSON)	Вибір або завантаження файлу з вхідними даними.
Аналіз (SWOT)	Checkbox Group (boolean)	Вибір складових: Сильні (Strengths), Слабкі (Weaknesses), Можливості (Opportunities), Загрози (Threats).
Модель (Parameters → Model)	Dropdown (enum)	Вибір підходу моделювання: Cognitive, Bayesian, Statistical, Machine Learning, Deep Learning, System Dynamics, Agent-based.
Алгоритм (Parameters → Algorithm)	Dropdown (enum)	Вибір алгоритму: Ймовірнісні та басові, Monte Carlo Simulation, Decision Tree, Regression Models тощо.
Weight (вага)	Numeric Slider (integer)	Встановлення ваги фактора у межах від 1 до 200.
Sweet (рівень інтенсивності)	Numeric Slider (integer)	Встановлення інтенсивності ознаки у межах від 1 до 100.
Вихідні параметри	Checkbox Group (boolean)	Вибір категорій результатів: Аналітичні індикатори, Метрики моделей, Бізнес-показники, Візуалізації.

Така композиція інтерфейсу дозволяє аналітику одночасно контролювати вихідні дані, параметри моделей і результати симуляцій, мінімізуючи перемикання між екранами. Інтуїтивні вкладки, геоприв'язана візуалізація та структурований репозиторій моделей утворюють цілісний робочий простір, що прискорює проведення SWOT/PEST-аналізів, сценарного прогнозування та експериментів *what-if*. У результаті користувач отримує зручний інструмент для швидкого формування обґрунтованих управлінських рішень.

Етап 5. Побудова моделей і прогнозів

Інтерфейс «Побудова моделей і прогнозів» концентрує увесь процес прогнозування на одній панелі: у центрі — інтерактивна схема $\text{Data} \rightarrow \text{Model} \rightarrow \text{Forecast}$ із fan-chart-графіком, праворуч — панель налаштувань горизонтів (1, 3, 6, 12 міс) та алгоритмів (ARIMA, ETS, Regression, Spline, Prophet), а під графіком — «simulation slider» для миттєвої прокрутки *what-if*-сценарію. Така компоновка дає змогу аналітику швидко під'єднувати дані, перемикати моделі й оцінювати невизначеність прогнозу в реальному часі.



Рисунок 5.9 – Вікно побудови моделей і прогнозів

Нижче стисло описано ключові елементи екрану «Побудова прогновної моделі»: центральна діаграма *Дані* → *Модель* → *Прогноз* із fan-chart для візуальної перевірки результатів, права панель налаштувань горизонту та алгоритму прогнозування, а також слайдер симуляції для миттєвого перегляду *what-if*-сценаріїв.

Таблиця 5.7 - Опис процесів реалізації побудови когнітивних моделей

Блок	Компоненти	Що робить користувач
Forecast diagram (центр)	• Схема «Дані → Модель → Прогноз» • Лінійний графік із довірчим інтервалом (fan-chart)	Візуально перевіряє результати моделі, оцінює різницю між прогнозом і фактичними даними.
Settings (праворуч)	• <Radio> Горизонт прогнозу — короткостроковий, середньостроковий, довгостроковий • <Checkbox List> Моделі прогнозування — SARIMA, VAR, RNN, LSTM, AutoML • <Slider> Довірчий інтервал — від 0% до 100%	Обирає часовий горизонт прогнозу, алгоритм або кілька моделей для порівняння, налаштовує рівень довірчого інтервалу.
Model metrics (нижня панель)	• Кнопки для відображення RMSE, MAE, MAPE, R^2 • Опція «Вивести у таблицю»	Переглядає точність моделі, експортує метрики у табличний формат.
Export & Save (нижня панель праворуч)	• Кнопка «Зберегти файл SCV/XLS» • Кнопка «Зберегти до репозиторію» • Кнопка «Зберегти до файлу»	Експортує результати у файл, зберігає модель у репозиторій для повторного використання.
Action (внизу праворуч)	• Кнопка «Продовжити»	Переходить до наступного етапу процесу прогнозування.

Поєднання наочного fan-chart-графіка, гнучких налаштувань горизонту та алгоритму, а також інтерактивного слайдера створює зручний робочий простір,

де результати моделювання одразу відображаються у візуальній формі. Це скорочує час на тестування гіпотез, покращує прозорість прогностичних припущень і допомагає швидше ухвалювати управлінські рішення, спираючись на бренч-сценарії та оцінку ризиків.

5.3 Реалізація рекомендаційної підсистеми у процесі прийняття рішень в управлінні регіональним розвитком

Актуальною є задача розроблення інформаційних технологій, моделей та методів накопичення великих обсягів інформації про поведінку користувачів мережі Інтернет, потенційні загрози, події рішення, ситуації. Важливо накопичити практичний вирішення задач та реагування на проблемні ситуації, сформулювати рекомендації щодо дій при виборі варіантів побудови прогнозів, сценаріїв, тощо. Одним із шляхів розв'язання вказаних задач є розроблення спеціалізованих рекомендаційних систем. Їх використання є звичайною практикою у багатьох сферах [66-68]. Рекомендаційні системи на основі колаборативної фільтрації, фільтрації на основі контенту і гібридні використовуються досить давно в торгівлі та маркетингу, освіті, тощо. Розвиток штучного інтелекту дозволив подолати основні недоліки традиційних рекомендаційних систем: проблему «холодного запуску», масштабування, обмежень стосовно персоналізації, відсутність різноманіття, розріженість даних, проблеми інтерпретації, і що особливо корисно, використовувати у рекомендаційних системах динамічні та розмовні взаємодії, мультимодальні рекомендації та контекстну обізнаність, тощо.

Рекомендаційні системи можуть персоналізувати інформаційну підтримку, аналізуючи попередній досвід працівників та аналогічні ситуації, сценарії використання рекомендацій, які були результативними у аналогічних випадках. Такі системи здатні швидко аналізувати великі обсяги даних з різних джерел, включаючи сенсори, історичні записи та технічну документацію, щоб надавати

оптимальні сценарії реагування. Оскільки оперативність прийняття рішень є критичним фактором в управлінні, особливо, кризовому. Рекомендаційні системи допомагають мінімізувати затримки завдяки обробці та синтезу ключової інформації в режимі реального часу.

Основними компонентами такої системи є процеси генерації, оцінювання та коригування рекомендацій. Генерація кандидатів передбачає формування набору потенційних рекомендацій шляхом аналізу даних про попередні події, поведінку користувачів та історичні сценарії. Після цього відбувається оцінювання кандидатів, яке здійснюється за допомогою алгоритмів ранжування, що дозволяють відібрати найбільш релевантні рекомендації відповідно до контексту конкретної ситуації. На завершальному етапі коригування рейтингу рекомендації адаптуються з урахуванням актуальних умов, таких як час, місцезнаходження та пріоритетність варіантів рішення, що забезпечує їхню максимальну ефективність.

Основні принципи даної рекомендаційної системи полягають у тому, що система аналізує характеристики об'єктів, таких як текст, метадані або мультимедійні елементи, і на основі цього формує рекомендації. Вона створює профіль користувача, враховуючи його попередні вподобання та взаємодії, порівнює ці дані з характеристиками доступних об'єктів, щоб визначити рекомендації, що найбільше відповідають досліджуваній ситуації.

Традиційні системи рекомендацій, такі як контентна фільтрація та колаборативна фільтрація, базуються на аналізі історичних взаємодій та явних атрибутів користувачів і об'єктів. Хоча ці підходи успішно застосовуються у багатьох сценаріях, вони мають певні обмеження. Проблема холодного старту, коли нові користувачі чи об'єкти не мають достатньої історичної інформації, а також проблема розрідженості даних, коли величезна кількість елементів призводить до мінімальної кількості взаємодій на одного користувача, можуть суттєво знижувати релевантність рекомендацій. Саме тому виникає потреба у більш потужних методах, здатних виявляти складні латентні закономірності з різномірних джерел даних. Методи штучного інтелекту, зокрема глибоке

навчання, стають ефективною альтернативою, адже вони забезпечують підвищену адаптивність, можливість інтегрувати різноманітні типи даних та покращену точність прогнозування в умовах швидкоплинних сценаріїв. Тому, затребуваними стають альтернативні методи формування рекомендацій, зокрема, нейронні мережі. Конволюційні нейронні мережі (CNN) підходять для обробки неструктурованих мультимедійних даних завдяки їх здатності ефективно видобувати ознаки з зображень, тексту, аудіо та відео. Використання CNN дозволяє не тільки покращити якість видобутку інформації з візуального контенту, але й усунути проблему холодного старту або доповнити традиційні методи, такі як колаборативна фільтрація. Крім того, CNN можна застосовувати для роботи з неевклідовими даними (непорядковими або ієрархічними), наприклад, у соціальних мережах, мережах взаємодії білка або графах знань. Рекурентні нейронні мережі (RNN) можуть також бути використані для обробки послідовних даних завдяки здатності визначати часову динаміку взаємодій і послідовних патернів поведінки користувачів. RNN дозволяє моделювати часову динаміку взаємодій користувачів та виявляти послідовні патерни їхньої поведінки, що особливо важливо для створення релевантних рекомендацій навіть при короткострокових сесіях без довгострокових даних про користувача. За допомогою RNN можна створювати рекомендації на основі сесійних даних, навіть без ідентифікації користувача, виходячи з історії кліків. Це дозволяє значно підвищити релевантність рекомендацій навіть у випадку короткострокової взаємодії з системою. Інший варіант – автоенкодер. За своєю сутністю, автоенкодер є нейронною мережею, що відновлює вхідні дані на вихідному шарі. Він містить внутрішній прихований шар, який формує код для представлення вхідної інформації. Автоенкодер складається з двох основних частин: енкодера, який перетворює вхідні дані у компактне представлення (код), та декодера, що відтворює вихід, базуючись на цьому коді. Основна перевага автоенкодера полягає в його здатності знижувати розмірність даних, реконструювати їх та видобувати важливі ознаки.

Гібридні рекомендаційні системи поєднують сильні сторони класичних методів (як-от контентна та колаборативна фільтрація) із потужними алгоритмами штучного інтелекту, наприклад, глибокими нейронними мережами. Такий підхід дозволяє швидко формувати початковий набір кандидатів за допомогою прозорих і інтерпретованих методів, а потім уточнювати їх рейтинг, використовуючи можливості моделювання складних нелінійних залежностей і інтеграції різнорідних даних. Це особливо важливо у сценаріях, де критична точність і адаптивність рекомендацій є запорукою ефективного прийняття рішень, наприклад, у ситуації військового конфлікту, медицині, критичної інфраструктурі або кібербезпеці, де швидкість реагування і висока релевантність прогнозів мають вирішальне значення. Застосування Deep Q-Network (DQN) у рекомендаційних системах такого типу набуває особливого значення. Новизна підходу з використанням DQN полягає у поєднанні методів глибокого навчання та reinforcement learning для створення адаптивної рекомендаційної системи, що самонавчається. Її основними перевагами є:

- адаптивність в режимі реального часу (дозволяє системі безперервно оновлювати свої стратегії на основі поточного стану та зворотного зв'язку від користувачів, що забезпечує високу реактивність у змінних умовах кібербезпеки);
- оптимізація кумулятивної винагороди (завдяки максимізації кумулятивного очікуваного винагородження, система може ефективно прогнозувати не тільки короткострокові, але й довгострокові наслідки прийнятих дій, що є критичним для забезпечення стійкості критичної інфраструктури);
- інтеграція з гібридними системами (може бути інтегроване з традиційними методами рекомендацій (наприклад, контентною або колаборативною фільтрацією), що дозволяє отримати гібридний підхід з високою точністю прогнозів і адаптивністю);
- новий рівень автоматизації прийняття рішень (забезпечує можливість розробки систем, які автоматично приймають рішення на основі історичних

даних і поточного контексту, зменшуючи потребу в «ручному» втручанні та підвищуючи ефективність управління загрозами).

Таким чином, використання DQN у рекомендаційних системах є новим підходом, що сприяє підвищенню адаптивності, ефективності та автоматизації систем управління, відповідаючи на сучасні виклики.

Отже, гібридний підхід до побудови рекомендаційних систем, який поєднує класичні методи (контентну та колаборативну фільтрацію) із сучасними моделями штучного інтелекту (зокрема, глибокими нейронними мережами), забезпечує значне покращення точності прогнозування та адаптивності рекомендацій.

Виконане дослідження свідчить про те, що інтеграція латентного аналізу та видобутку ознак дозволить ефективно вирішувати проблеми холодного старту та розрідженості даних, що є важливою перевагою.

Результати дослідження «класичного» підходу до побудови рекомендаційних систем та за використання штучного інтелекту, представлені в табл. 5.8

На етапі проектування інформаційної технології, було проведено низку експериментів, які дозволили обрати оптимальний варіант рекомендаційної системи. Для порівняння було обрано класичні рекомендаційні системи (на основі контентної та колаборативної фільтрації) та системи, побудовані на базі штучного інтелекту (глибокі нейронні мережі, DNN). Дослідження виконано на наборі даних `cybersecurity_attacks.csv` (цільова змінна – «Action Taken»). Набір даних містить інформацію про кібератаки, де кожен приклад атак описується низкою ознак (наприклад, характеристики атаки, мережеві показники, часові параметри тощо), цільовою змінною є «Action Taken» – дії, які було виконано для реагування на атаку.

Даний набір даних використано як тестовий. Використані у дослідженні «класичні рекомендаційні системи» реалізують методи контентної та колаборативної фільтрації, які базуються на аналізі історичних взаємодій та явних атрибутів.

Таблиця 5.8 – Аналіз переваг та недоліків «класичних» рекомендаційних систем та систем, на основі штучного інтелекту

Характеристика	Класичні рекомендаційні системи	Системи на основі штучного інтелекту
Вхідні дані	Використовують явні атрибути користувача та об'єкта, історичні рейтинги чи взаємодії	Можуть інтегрувати різноманітні дані, що дає можливість побудови багатовимірних представлень
Проблема холодного старту	Є проблема «холодного» старту через недостатню історичну інформацію для нових користувачів або об'єктів	Завдяки використанню латентних представлень та вбудовування (embeddings) краще долають проблему холодного старту
Розрідженість даних	Можуть страждати від розрідженості даних, коли взаємодій користувачів з великою кількістю об'єктів замало для точного аналізу	Використання методів видобутку латентних ознак допомагає зменшити вплив розрідженості та покращити якість рекомендацій
Точність рекомендацій	Працюють добре при наявності достатніх історичних даних, обмежені здатністю моделювати складні взаємозв'язки між ознаками	Можуть моделювати нелінійні залежності та взаємодії між різними ознаками, що часто забезпечує вищу точність та релевантність рекомендацій
Адаптивність	Зазвичай мають обмежену адаптивність і потребують ручного оновлення моделей при зміні умов	Забезпечують енд-ту-енд навчання, оперативно адаптуються до змін у вподобаннях користувачів перенавчаючись на нових даних
Обчислювальна складність	Зазвичай мають нижчу обчислювальну складність і використовують прості алгоритми, що може бути перевагою для невеликих або середніх обсягів даних	Вимагають значних обчислювальних ресурсів через використання глибоких архітектур, але є технічні можливості масштабувати їх на великих даних
Інтерпретованість	Більш прозорі та інтерпретовані, оскільки базуються на явних співвідношеннях між даними	Часто є «чорним ящиком», проте сучасні методи explainable AI допомагають покращити інтерпретованість моделей.
Інтеграція різноманітних даних	Обмежені в здатності інтегрувати різні типи даних, що може впливати на якість рекомендацій у складних сценаріях	Можуть ефективно об'єднувати різні модальності даних (наприклад, зображення, текст, аудіо), для отримання кращого уявлення про користувача

Рекомендаційні системи на основі штучного інтелекту використовували штучні нейронні мережі. Для прогнозування дій у відповідь на кібератаку було побудовано глибоку нейронну мережу з прихованими шарами та вихідним шаром, що використовує функцію softmax для отримання розподілу ймовірностей для кожного класу дій. Для порівняння моделей використовувалися метрики: Accuracy (точність), Precision (точність прогнозування), Recall (повнота), F1 Score, середня квадратична помилка (RMSE). Також досліджувались характеристики, що стосуються «холодного» старту, адаптивності до нових даних, обчислювальної ефективності та інтерпретованості. Результати експериментів узагальнено в таблиці 5.9.

Таблиця 5.9 – Результати оцінювання ефективності рекомендаційних систем

Модель	Accuracy, %	Precision, %	Recall, %	F1 Score, %	RMSE
Контентна фільтрація	68	66	65	65.5	0.80
Колаборативна фільтрація	70	68	67	67.5	0.78
Гібридний класичний підхід	72	70	68	69	0.75

Як видно з таблиці 5.9, гібридний підхід, який поєднує переваги як контентної, так і колаборативної фільтрації, показує кращі результати за всіма метриками: вищу точність, кращу здатність знаходити релевантні рекомендації та меншу середню квадратичну помилку.

Отримані результати свідчать про те, що інтеграція різних класичних методів може покращити ефективність рекомендаційної системи, забезпечуючи більш точний та релевантний вибір пропозицій для користувачів.

У дослідженні також були проведені експерименти для порівняння точності рекомендацій за застосування різних підходів штучного інтелекту. Для цього використовувалися кілька моделей, зокрема глибокі нейронні мережі (DNN), рекурентні нейронні мережі (RNN) та автоенкодера, що дозволило розглянути як статичні, так і послідовні дані. Експерименти проводилися на

наборі даних про кібератаки, цільовою змінною була дія, яку виконували для реагування на атаку.

Основними метриками оцінки точності моделей обрані Accuracy (Точність) - відсоток правильних прогнозів від загальної кількості прикладів, Precision (Точність прогнозування) - показник, що характеризує правильність позитивних прогнозів, Recall (Чутливість) – здатність моделі знаходити всі позитивні приклади, F1 Score - гармонійне середнє між precision та recall; RMSE (середньоквадратична помилка) - показник відхилення прогнозів від реальних значень.

Набір даних був попередньо підготовлений та розбитий на навчальну, валідаційну та тестову вибірки. Моделі навчалися з використанням енд-ту-енд навчання, а їхні гіперпараметри оптимізувалися за допомогою пошуку по сітці (grid search) або адаптивних методів оптимізації (наприклад, Adam).

Після навчання результати моделей були порівняні за вищезазначеними метриками. Результати показали, що глибокі нейронні мережі (DNN) продемонстрували найкращу загальну точність і здатність правильно прогнозувати оптимальні дії у відповідь на кібербезпекові атаки.

Результати експериментів представлені у таблиці 5.10.

Таблиця 5.10 - Узагальнені результатів експериментів

Модель	Accuracy, %	Precision, %	Recall, %	F1 Score, %	RMSE
DNN	80	78	75	76	0.65
RNN	78	76	73	74.5	0.68
Autoencoder-based	75	73	70	71.5	0.72

Рекурентні нейронні мережі незначно поступалися DNN за всіма метриками, проте добре працювали з послідовними даними, що може бути корисно для аналізу сесійних взаємодій. Модель, побудована на автоенкодерах, показала дещо гірші результати, що, ймовірно, пов'язано з меншою здатністю виявляти складні нелінійні залежності порівняно з глибокими мережами. Таким

чином, експериментальним шляхом підтверджено ефективність застосування методів штучного інтелекту, зокрема DNN, для підвищення точності рекомендаційних систем, що є особливо важливим у критичних ситуація, де точність і швидкість реагування мають вирішальне значення.

Результати дослідження застосування гібридного підходу, який інтегрує переваги класичних методів (контентної, колаборативної фільтрації) та моделей на основі штучного інтелекту (наприклад, DNN), представлені у табл. 5.11. Цей підхід дозволяє поєднати швидкість та інтерпретованість традиційних алгоритмів з високою точністю та адаптивністю глибинних моделей. Результати експериментів демонструють, що гібридна система забезпечує покращення всіх основних метрик, таких як точність, precision, recall, F1 score, а також зниження середньої квадратичної помилки (RMSE).

Таблиця 5.11 – Результати дослідження використання гібридного підходу

Модель	Accuracy,%	Precision,%	Recall%	F1 Score,%	RMSE
Гібридний підхід	82	80	78	79	0.60

Як видно з таблиці 5.11, гібридний підхід дозволяє досягти вищих показників точності - 82%, проти ~72–80% у класичних системах та системах із використанням штучного інтелекту, підвищену точність прогнозування (precision), кращу здатність знаходити релевантні рекомендації (recall і F1 score), а також зменшує середню квадратичну помилку (RMSE). Це свідчить про ефективність комбінації класичних методів із сучасними моделями штучного інтелекту для побудови рекомендаційних систем, що особливо важливо у сценаріях, де потрібна висока точність і адаптивність рішень.

Проведені експерименти показали, що інтеграція класичних методів рекомендаційних систем із сучасними моделями штучного інтелекту дозволяє значно покращити якість прогнозування дій для критичної інфраструктури. Глибокі нейронні мережі (DNN) виявили здатність ефективно видобувати латентні ознаки та моделювати складні нелінійні залежності, що сприяє

підвищенню точності рекомендацій за основними метриками (Accuracy, Precision, Recall, F1 Score та RMSE). Проте, як показали результати, класичні системи, побудовані на контентній та колаборативній фільтрації, мають переваги в обчислювальній ефективності та інтерпретованості. Гібридний підхід дозволяє використати сильні сторони обох методів: швидке первинне формування кандидатів класичними алгоритмами та точне уточнення рейтингу за допомогою моделей штучного інтелекту. Отже, комбінований метод забезпечує вищу загальну продуктивність, що є критично важливим для оперативного прийняття рішень у кризових ситуаціях, але водночас потребує додаткових обчислювальних ресурсів. В подальшому \ можливість оптимізації гібридної системи, зокрема, шляхом адаптації вагових коефіцієнтів інтеграції та застосування сучасних методів explainable AI для підвищення прозорості моделей.

Оскільки, на даний час, існують часткові рішення, які не задовольняють вимогам щодо якості та оперативності підтримки прийняття рішень, зокрема, надання рекомендацій щодо дій по захисту об'єктів критичної інфраструктури, розміщення біженців, реагування на екологічні небезпеки та техногенні аварії.. Загальна характеристика роботи розробленої системи може бути представлена у три етапи.

На етапі генерація кандидатів система відбирає підмножину кандидатів із великого набору елементів. Для цього використовуються декілька генераторів, кожен з яких формує свою підмножину, виконуючи фільтрацію на основі: контенту, онтологій, знань або моделей глибоких нейронних мереж (наприклад, модель Softmax). Далі виконується оцінювання і ранжування (скорінг). З множини елементів вибирають елементи для відображення користувачеві, використовуючи більш точні моделі. До таких моделей належать: градієнтний бустинг, нейронний колаборативний фільтр, моделі нечіткої логіки. Далі виконується зміна рейтингу (переоцінка). На цьому етапі система враховує додаткові фактори, такі як: виключення непопулярних елементів; підвищення оцінки новішого контенту для забезпечення різноманітності. Для вирішення цих

завдань застосовані рекурентні нейронні мережі, трансформери, а також рангові моделі. Особливістю даної рекомендаційної системи є використання штучного інтелекту на всіх етапах генерування рекомендацій.

Методика використання математичних моделей у рекомендаційній системі представлена на рис. 5.12.

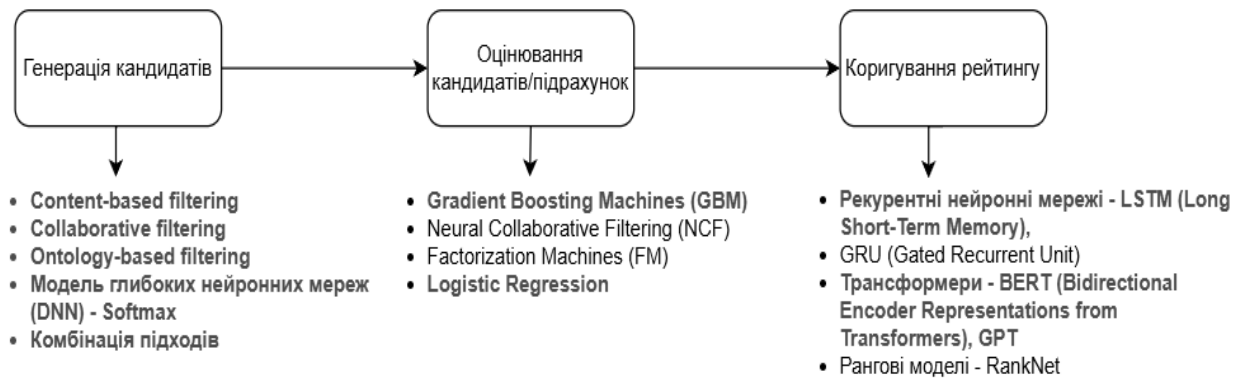


Рисунок 5.12 - Методика використання математичних моделей у рекомендаційній системі

Як видно з рис. 5.12, на етапі генерації кандидатів система відбирає підмножину кандидатів із великого набору елементів. Для цього використовуються декілька генераторів, кожен з яких формує свою підмножину, виконуючи фільтрацію на основі контенту, онтологій, знань, або моделі глибоких нейронних мереж (наприклад модель Softmax). На етапі оцінки і ранжування (скорінгу) з множини елементів вибирають елементи для відображення, використовуючи більш точні моделі, такі, як градієнтний бустинг, нейронний колаборативний фільтр або моделі нечіткої логіки. На етапі зміни рейтингу (переоцінки) система враховує додаткові фактори, такі як виключення непопулярних елементів або підвищення оцінки новішого контенту, щоб забезпечити різноманітність і неупередженість. На даному етапі можуть бути застосовані рекурентні нейронні мережі або трансформери, а також рангові моделі.

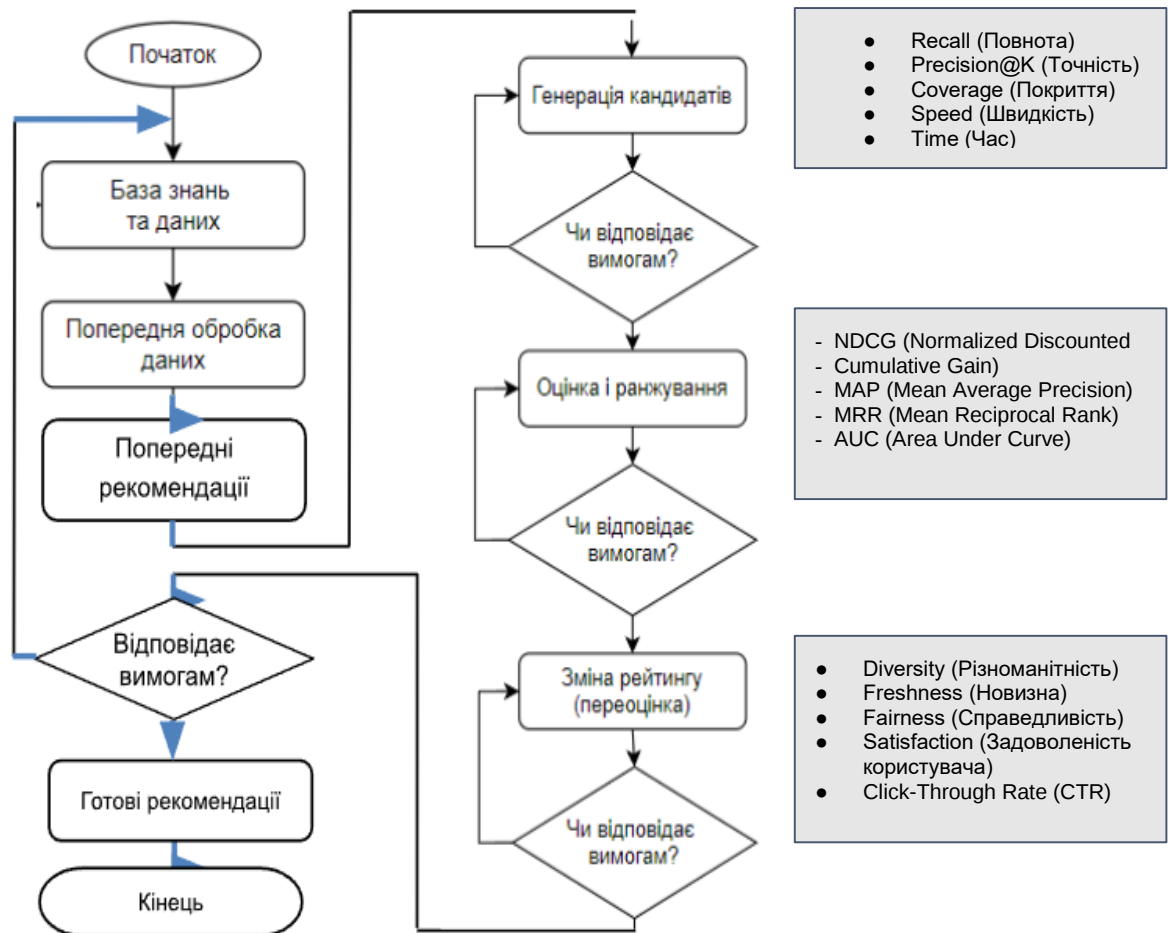


Рисунок 5.13 - Алгоритм роботи рекомендаційної системи

Перевірка працездатності розробки виконана на наборі даних, представлених на <https://www.kaggle.com/datasets/sandeepkumar69/website-traffic-data> [70].

Математичні моделі, реалізовані у рекомендаційній системі, охоплюють всі етапи обробки даних, необхідні для формулювання рекомендацій. Однією з особливостей даного підходу є методика виконання попередньої обробки даних, виявлення та аналіз аномалій. Заповнення пропусків виконується залежно від того, наскільки вагомою є частка пропущених значень. Якщо пропущено більше 20% значень, то для дослідження процесу застосовується методика визначення подібності часових рядів. Для виключення впливу випадкової складової, застосовується метод згладжування часових рядів.

На наборах даних, використаних для перевірки працездатності рекомендаційної системи опрацьовані всі етапи розробленого алгоритму роботи рекомендаційної системи (рис. 5.14).



Рисунок 5.14 – Методика опрацювання даних про кібератаки у пропонованій системі

Для попередньої обробки даних було використано метод ковзного середнього, який полягає у звичайній заміні значень членів ряду на середнє арифметичне за відповідним вікном.

$$\bar{X}(k) = \frac{1}{n} \sum_{t=k}^{n+k} X(t), \quad (5.1)$$

де n - розмір вікна (період згладжування), k - номер члена ряду, значення якого замінюється середнім

Експоненційний фільтр або експоненційне згладжування - це найпростіший і найпоширеніший фільтр, який використовується в багатьох

системах. Він так само називається фільтром першого порядку. Аналітично він може бути описаний таким рівнянням:

$$y(k) = a * y(k-1) + (1 - a) * x(k), \quad (5.2)$$

де $x(k)$ — введене значення на кроці k , $y(k)$ - відфільтровані значення на кроці k , а a є константою $(0;1)$, зазвичай між 0,8 та 0,99 "константа згладжування".

Перевірка на стаціонарність - тест Дікі-Фуллера показала, ADF Statistic: - 2.90, p-value: 0.045. Оскільки p-value менша за 0.05, можна припустити, що ряд є стаціонарним на рівні значущості 5%

У таблицях 5.12-8.14 представлені результати фільтрації даних, використані при проведенні дослідження.

Таблиця 5.12 – Результати експерименту (без фільтрації) за розподілу вибірки у співвідношенні 80/20

Modeltype	Model quality			Forecast quality			
	R ²	$\sum e^2(k)$	DW	MSE	MAE	MAPE	Theil
AR(1)	0,99	25644,67	2,15	49,82	41,356	8,37	0,046
AR(1,4)	0,99	25588,10	2,18	49,14	40,355	8,12	0,046
AR(1) + 1st order trend	0,99	25391,39	2,13	34,39	25,109	4,55	0,032
AR(1) + 4th order trend	0,99	25173,74	2,12	25,92	17,686	3,19	0,024

Таблиця 5.13 – Результати експерименту за розподілу вибірки у співвідношенні 80/20, використано експоненціальне згладжування (alpha = 0.7)

	Model quality			Forecast quality			
	R ²	$\sum e^2(k)$	DW	MSE	MAE	MAPE	Theil
AR(1)	0,99	24335,12	2,13	50,08	40,12	7,98	0,040
AR(1,4)	0,99	24453,1	2,12	47,15	39,56	7,68	0,042
AR(1) + 1 st ordertrend	0,99	25061,08	2,10	34,07	23,26	3,55	0,033
AR(1) + 4 th ordertrend	0,99	23881,14	2,07	24,1	16,58	3,05	0,022

Таблиця 5.14 - Результати експерименту за розподілу вибірки у співвідношенні 80/20, використано метод ковзного середнього (3 кроки)

Modeltype	Modelquality			Forecastquality			
	R ²	$\Sigma e^2(k)$	DW	MSE	MAE	MAPE	Theil
AR(1)	0,99	23355,54	2,10	44,80	39,65	7,44	0,036
AR(1,4)	0,99	24132,15	2,08	46,45	38,34	6,78	0,034
AR(1) + 1 st ordertrend	0,99	23861,65	2,07	31,07	22,08	3,12	0,027
AR(1) + 4 th ordertrend	0,99	21887,54	2,05	21,24	13,13	2,55	0,017

Отже на наборах даних щодо кібератак, було проведено експеримент із використанням можливості застосування рекомендаційних систем. В рамках проведення експерименту було виконано:

- розроблення та тестування прототипу рекомендаційної системи, адаптованої до специфічних потреб та особливостей предметної області, а саме:

- застосування методів збору та аналізу даних для підвищення точності рекомендацій.
- оцінювання ефективності рекомендаційної системи через серію експериментів та тестових сценаріїв.

- застосування методів штучного інтелекту для побудови рекомендацій, а саме:

- використання алгоритмів машинного навчання та глибокого навчання для автоматичного визначення та прогнозування потреб користувачів.
- оптимізація моделей на основі результатів експериментів з метою покращення релевантності та точності рекомендацій.

- впровадження поведінкових показників для аналізу та надання рекомендацій, а саме:

- збір даних про поведінкові патерни користувачів для персоналізації рекомендацій.
- аналіз взаємодії користувачів із системою для виявлення закономірностей.

- налаштування системи на основі поведінкових показників для забезпечення індивідуального підходу в рекомендаціях та підвищення задоволеності користувачів.

Таким чином, використання засобів штучного інтелекту дозволило швидко виявляти аномальні або недостовірні дані про стан та функціонування об'єкту, що забезпечує раннє виявлення загроз, вчасне реагування на них, користуючись рекомендаціями системи, що знизить вплив людського фактору та підвищить ефективність підтримки прийняття рішень у системі управління регіональним розвитком. В подальшому пропонується рекомендаційна система може бути удосконалена за рахунок розширення набору використовуваних математичних моделей та текстової аналітики.

Висновки до розділу 5

1. Представлено методологію побудови інформаційних технологій для підтримки прийняття рішень у системі управління регіональним розвитком.
2. Виходячи з особливостей соціально-економічної систем регіонального рівня та наявної організації структури управління регіонами та громадами, просторового розташування складових, розроблювані інформаційні технології поєднано у інформаційно-керуючій системі, побудованій за мережецентричним принципом, взаємодія об'єктів та процесів управління якої відбувається у єдиному інформаційному просторі.
3. Розроблено метод синтезу інформаційних технологій на основі багатомодельного й багатокритеріального підходів, який, на відміну від відомих, передбачає системне використання методів збору, оброблення, інтегрування та аналізу різнотипної інформації, моделювання, прогнозування і багатокритеріального прийняття рішень, із можливістю автоматизованого інтегрування даних, адаптивним налаштуванням та

урахуванням досліджуваних процесів, що збільшить ефективність комплексного аналізу.

4. Представлено прототип розробленого програмного забезпечення, наведено приклади його використання.
5. Розроблено методику автоматизованого реагування на загрози різної природи, яка на відміну від чинних, реалізує вибір оптимальної стратегії з множини можливих, з урахуванням прогностичної інформації та обмежень ресурсів, що дозволяє мінімізувати втрати та підвищити стійкість системи.
6. Визначено, що запропонована СППР відрізняється застосуванням мережецентричного підходу до побудови і має гнучку архітектуру, що є важливим фактором для модифікації розроблених методів і методик аналізу даних, а також продовження виконання досліджень в даній області.
7. Розроблені та реалізовані методи призначені насамперед для автоматизації процесу підтримки прийняття рішень в управлінні регіональним розвитком в умовах невизначеності та необхідності швидкого реагування на загрози.

РОЗДІЛ 6

ПРИКЛАДИ ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В УПРАВЛІННІ РЕГІОНАЛЬНИМ РОЗВИТКОМ

6.1 Інформаційна технологія формування місцевого бюджету

В даному розділі представлені приклади реалізації розробленої методології побудови інформаційних технологій для підтримки прийняття рішень у системі управління регіональним розвитком для основних сфер економіки.

Як було встановлено в попередніх розділах для розробки сценаріїв розвитку соціально-економічних систем доцільно використовувати методику, яка передбачає поєднання різних засобів моделювання.

Використання когнітивного моделювання (як побудова карт, так і побудова матриць) та ймовірнісного моделювання у процесі формування сценаріїв розвитку соціально-економічних систем передбачає виконання наступних кроків:

КРОК 1. Визначення цільових установок побудови сценаріїв.

КРОК 2. Визначення основних факторів, що впливають на розвиток системи.

КРОК 3. Відбір із набору факторів найбільш значимих.

КРОК 4. Визначення сильних та слабких характеристик системи на основі відібраних найбільш важливих змінних.

КРОК 5. Формування множини вхідних концептів для побудови когнітивної карти.

КРОК 6. Побудова когнітивної карти на основі відібраних на кроці 5 концептів.

КРОК 7. Аналіз когнітивної карти.

КРОК 8. Оцінювання системного ризику.

КРОК 9. Формулювання можливих сценаріїв розвитку подій.

КРОК 10. Оцінювання якості побудованих сценаріїв, обґрунтування вибору найкращого та найімовірнішого.

Дану методику було застосовано для побудови сценаріїв розвитку

соціально-економічної системи – агропромислового комплексу регіону. Для дослідження використано такі фактори, як природно-кліматичні та екологічні умови, ресурсний потенціал, соціально-економічний розвиток, ринкова інфраструктура тощо. Всього, було опрацьовано набори вхідних даних, що складаються із 56 показників соціально-економічного розвитку України та Черкаської області за 2010–2013 рр., а окремі показники і за 2003–2013 рр.

Для обґрунтування цільових установок побудови сценаріїв було використано не лише статистичні дані та висновки експертів, а й виконано аналіз Інтернет-джерел засобами SAS Textual Analytics. Для аналізу корпусу текстів було використано SAS Textual Analytics Suit, що складається з SAS Content Categorization та SAS Sentimental Analysis Studio за допомогою яких була розроблена таксономія із 604 правил, опрацьовано більше 180 Інтернет-джерел, включаючи офіційні джерела органів державного управління, місцевого самоврядування, приватні сайти, сайти політичних партій, блоги.

В результаті дослідження виявлено, що чинники, які впливають на розвиток соціально-економічної системи регіону необхідно розглядати не лише з точки зору загальної характеристики окремого фактору, а й в контексті конкретної системи та її цільового спрямування. Наприклад, показник рівня безробіття (як вважає значна частина населення – 59% згідно дослідження Інтернет-ресурсів є негативним чинником розвитку регіону) насправді, як показали розрахунки, не є негативним фактором, оскільки загальна кон'юнктура ринку праці вносить свої корективи. Скоріше, негативним фактором є вікова структура безробіття, зниження економічної активності населення та частка прихованого безробіття, які призводять до зниження ємності споживчого ринку. Розвиток інвестиційної діяльності в агропромисловому виробництві є фактором впливу на активізацію розвитку економіки та підвищення соціально-економічного рівня населення, з другого – зростання економіки та покращення життя населення сприяють притоку інвестиційних ресурсів, адже основним інвестором у сільській місцевості є жителі регіону. Це, як показало дослідження, є особливістю соціально-економічної системи регіону, в економіці якого переважає

агропромислове виробництво.

При виборі цільової установки було враховано двонаправленість факторів впливу на розвиток соціально-економічної системи регіону та визначено, що цільовою установкою є розвиток агропромислового виробництва в регіоні.

Найбільш значимі фактори, що впливають на розвиток агропромислового виробництва регіону згруповані у табл. 6.1.

Таблиця 6.1. -- Групування факторів, що впливають на розвиток агропромислового виробництва регіону

Групи факторів	Характер впливу	Ступінь впливу
Політичні		
часта зміна законодавчої бази, що регулює економічну діяльність в агропромисловому виробництві	негативний	висока
наявність бюрократичних бар'єрів, відсутність координації дій органів влади різних рівнів та місцевого самоврядування	негативний	помірна
державна підтримка	позитивний	помірний
Економічні		
зростання відкритості національної економіки та конкуренції на світових ринках продовольства та сільськогосподарської сировини	негативний	помірна
нестабільність фінансово-кредитної системи, обмеженість обігових коштів для ведення бізнесу	негативний	висока
підтримка пріоритетних проектів в АПК, надання податкових та інших пільг інвесторам, розвиток державно-приватного партнерства	позитивний	помірна
значна інвестиційна привабливість окремих галузей (підприємств) агропромислового виробництва	позитивний	висока
євроінтеграція та лібералізація торгівлі	позитивний	помірна
Соціальні		
наявність дешевих та кваліфікованих трудових ресурсів	позитивний	помірна
зниження доходів населення	негативний	помірна
розвинені аграрна освіта та наука	позитивний	висока
Технологічні		
значний виробничий потенціал агропромислового комплексу	позитивний	висока
впровадження інноваційних технологій в харчовій промисловості та сільському господарстві	позитивний	висока
скорочення виробництва вітчизняної сільськогосподарської техніки та обладнання для агропромислового виробництва	негативний	висока

Сильні та слабкі характеристики системи узагальнені за допомогою SWOT-аналізу (табл. 6.2).

Таблиця 6.2. - Результати SWOT-аналізу розвитку соціально-економічної системи регіону

Позитивні фактори	Негативні фактори
Природно-кліматичні умови, природні ресурси та екологічна ситуація:	
- сприятлива екологічна ситуація; - наявність мінерально-сировинних ресурсів; - сприятливі метеорологічні умови; - наявність водосховищ для риборозведення; - більша частина земель області – чорноземи.	- викиди малоочищених стічних вод в навколишнє середовище; - прискорення процесів водної ерозії; - зниження залісненості територій; - значний рівень розораності; - недостатнє використання наявних природних ресурсів та кліматичних умов.
Демографічна характеристика та стан ринку праці:	
- стала структура трудового потенціалу, значна кількість населення працездатного віку; - значна частка кваліфікованих трудових ресурсів; - зростання чисельності самозайнятого населення; - скорочення прихованого безробіття; - підвищення попиту на робітничі професії; - скорочення ручної праці в сільському господарстві.	- скорочення чисельності населення, зокрема народжуваності; - старіння населення; - значне навантаження на одне вільне робоче місце; - низький рівень зайнятості; - низький рівень оплати праці; - рівень кваліфікації працівника практично впливає на можливість працевлаштування.
Економічний розвиток та галузева структура:	
- стала динаміка розвитку економіки – та галузева структура; - створення кооперативів та інших інтегрованих формувань; - стає зростання експортного потенціалу; - лідируючі позиції сільського господарства; - потужна хімічна промисловість; - зростання інвестиційної привабливості; - розвиток насінництва та племінного тваринництва; - розвиток переробної та харчової промисловості.	- згортання виробництва без урахування регіональних потреб; - зміни у галузевій структурі некеровані та стримують розвиток економіки області - повільно відновлюються міжгалузеві зв'язки в агропромисловому виробництві; - некерована інвестиційна діяльність призводить до хаотичних змін у галузевій структурі АПК; - переважання сировинної складової у експорті продукції АПК.
Ринкова інфраструктура:	
в області функціонує товарна біржа	низькі обсяги угод
Рівень життя населення:	
- зростання середньої заробітної плати; - споживання продуктів харчування згідно раціональних норм споживання; - зростання сукупних доходів та заощаджень домогосподарств; - зростання середнього розміру пенсії.	- випередження зростання рівня заробітної плати зростанням споживчих цін; - низька купівельна спроможність населення; - наявність значної диференціації між містами та між районами області; - зростання частки витрат на продукти харчування у домогосподарствах населення.

Подальший відбір факторів у даній задачі виконано за допомогою дворівневого багатокритеріального аналізу, оскільки має місце значна кількість різноспрямованих факторів, що впливають на розвиток системи.

Альтернативами першого рівня $A = \{a_1, a_2, \dots, a_n\}$ є виробництво продукції в окремих районах області. Множина альтернатив першого рівня оцінюється на множині критеріїв $C = C_1 \cup D$. Альтернативами другого рівня $B = \{b_{i1}, b_{i2}, \dots, b_{ik}\}$ є види продукції, що будуть реалізовані, обсяги реалізації та обсяги виробництва продукції підприємств агропромислового комплексу. Підмножина критеріїв $C_1 = \{c_1, c_2, \dots, c_p\}$, використовується для оцінки альтернатив виключно першого рівня, а множина критеріїв D – як для першого, так і для другого рівнів. Кожному елементу множини A («району») ставиться у відповідність множина альтернатив другого рівня (обсяги реалізації кожного виду сільськогосподарської продукції за окремим напрямком). Необхідно ранжувати альтернативи множини A враховуючи багатокритеріальні оцінки альтернатив множини B . Складністю даної задачі є те, що на першому рівні для вибору альтернатив використовуються переважно якісні ознаки, а на другому – кількісні. Тому наступним кроком є побудова когнітивної карти. В даній задачі побудовано нечітку когнітивну карту з використанням програмного забезпечення Mental Modeler. Перевагою нечітких когнітивних карт є їх наочність та зручність подання причинно-наслідкових зв'язків між концептами. Цільовим параметром когнітивної карти визначено зростання виробництва продукції АПК. Карта представлена на рис. 6.1.

Для аналізу когнітивної карти використано показники: консонансу (c_i) [6]:

$$c_{ij} = \frac{|v_{ij} + \overline{v_{ij}}|}{|v_{ij}| + |\overline{v_{ij}}|} \quad (6.1)$$

де v_{ij} , $\overline{v_{ij}}$ – пара зв'язків у транзитивно замкненій когнітивній матриці; дисонансу (d_i) [6],

$$d_{ij} = 1 - c_{ij} \quad (6.2)$$

впливу системи на концепт ($\overline{P_i}$) [6]

$$\overrightarrow{P_i} = \frac{1}{n} \sum_{j=1}^n p_{ij} \quad (6.3)$$

та концепту на систему ($\bar{P}_j$) [6]

$$\bar{P}_j = \frac{1}{n} \sum_{i=1}^n d_{ij} \quad (6.4)$$

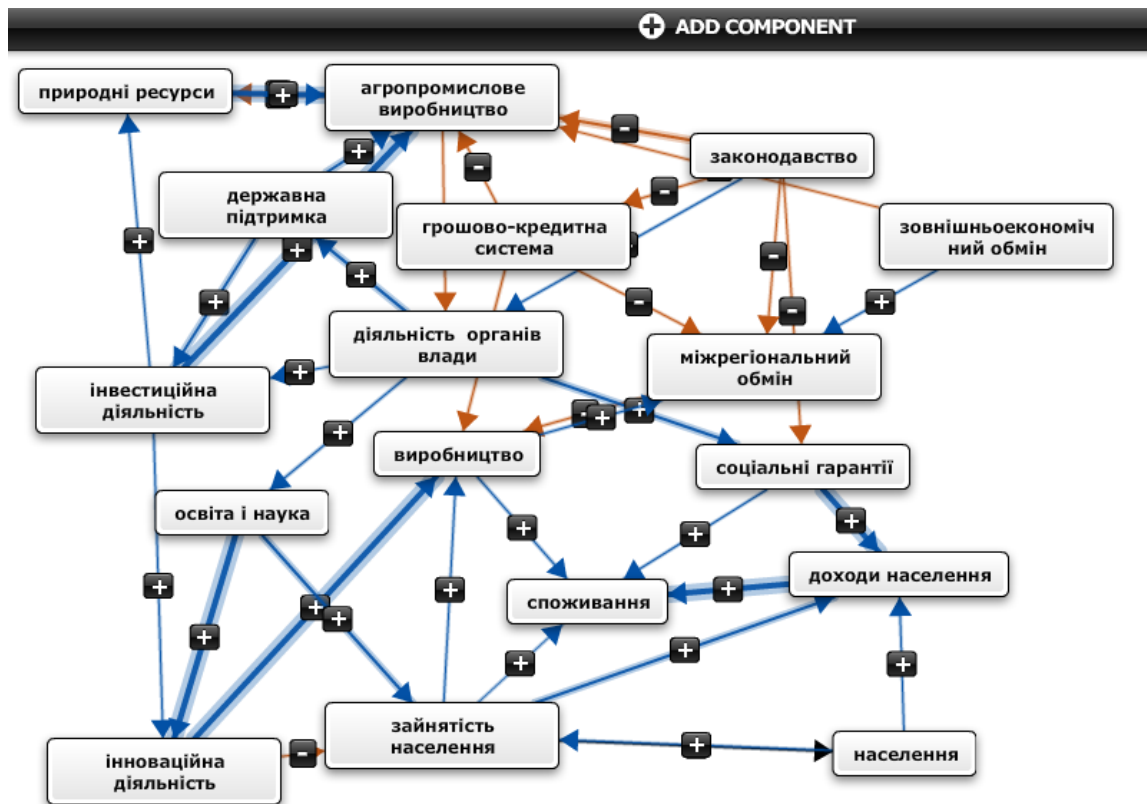


Рисунок 6.1.- Когнітивна карта розвитку агропромислового виробництва в регіоні

Показники консонансу впливу для більшості концептів є достатньо високими і знаходяться в межах від 0,64 до 0,91. Максимальний показник дисонансу достатньо низький – його значення становить 0,3. Тобто між відібраними концептами та системою існує суттєвий взаємовплив і вони можуть бути використані при побудові сценаріїв розвитку системи.

Для оцінювання системного ризику використано ймовірнісне моделювання за допомогою мережі Байєса. На основі побудованої топології мережі Байєса, визначаються найбільш значущі змінні, що впливають на цільову. Після цього будується рівняння множинної регресії із примусовим включенням до моделі виявлених змінних. Оцінка параметрів моделі виконується на основі рекурсивного методу найменших квадратів.

На основі одержаних результатів аналізу, сформовано наступні можливі

сценарії соціально-економічного розвитку регіону (табл. 6.3):

Таблиця 6.3. Розроблені сценарії розвитку агропромислового виробництва

Сценарій	Зміст сценарію
S1: Ефективне використання ресурсів	Даний сценарій передбачає значне надходження інвестиційних ресурсів в агропромислове виробництво за всіх джерел, в тому числі прямих іноземних інвестицій. В той же час наповнюваність місцевих бюджетів та розширені повноваження органів державного управління регіону будуть сприяти ефективному розвитку системи інвестиційної діяльності. Обсяги інвестицій, спрямовувані в агропромислове виробництво достатні для відтворення, управління економікою дієве, повноваження розподілені пропорційно, регіональний рівень державного управління має повноваження, що забезпечують автономність прийняття інвестиційних рішень. Сценарій можна вважати стійким, адже власні кошти підприємств тривалий час становлять основу їх інвестиційного забезпечення.
S2: Залежність від центру	За даного сценарію передбачається, що агропромислове виробництво буде забезпечене необхідними обсягами інвестиційних ресурсів, підвищиться наповнюваність місцевих бюджетів за рахунок ефективної роботи підприємств, сприятливому інвестиційному клімату, що буде приваблювати іноземних інвесторів, але є ризик надмірної централізації управління оскільки наповнюваність місцевих бюджетів є обмеженою за рахунок витрат на соціальний захист (низька заробітна плата в сільському господарстві та демографічне навантаження не дозволить значно збільшити надходження до місцевих бюджетів податку на доходи фізичних осіб), керування економікою, централізоване, що може продовжити сучасні тенденції.
S3: Диверсифікація ресурсів	Даний сценарій передбачає що інвестиційне забезпечення залишиться практично на існуючому рівні, але в країні буде впроваджено заплановані реформи у регіональному управлінні, підвищено автономію місцевого самоврядування. Виконання даного сценарію передбачає, що в умовах нестачі інвестиційних ресурсів для розвитку агропромислового виробництва, суб'єкти господарювання за підтримки центральних органів виконавчої влади та органів виконавчої влади в регіонах за підтримки місцевого самоврядування розпочнуть активну діяльність щодо пошуку додаткових джерел інвестиційних ресурсів та оптимізації використання внутрішніх інвестиційних ресурсів на рівні регіонів, активніше сприяти залученню зовнішніх, зокрема іноземних інвестицій.
S4: Централізація ресурсів	Сценарій передбачає суто централізоване управління, коли питання спрямування інвестиційних ресурсів в галузі національної економіки буде вирішуватися на загальнодержавному рівні, і пошук альтернативних варіантів їх залучення значно ускладниться в зв'язку з обмеженістю бюджетного фінансування та зниженням загальної інвестиційної привабливості, оскільки залежність від прийнятих на вищому рівні управління рішень буде дуже високою, держава не дбатиме про підтримку ініціативи регіонального рівня щодо оптимізації використання інвестиційних ресурсів адаптовано до умов регіонів.

Оцінювання якості побудованих сценаріїв, обґрунтування вибору найкращого та найімовірнішого виконано з використанням ймовірнісного моделювання.

В результаті дослідження ймовірності розроблених сценаріїв з використанням мережі Байєса, визначено, що найбільш ймовірним є (S3), результуючим показником для побудови мережі є валовий внутрішній продукт у фактичних цінах, млн грн та набір із 26 соціально-економічних показників України: x_{01} – валове заощадження млн грн; x_{02} – витрати населення, млн грн; x_{03} – доходи населення, млн грн; x_{04} – рівень безробіття (за методологією МОП), %; x_{05} – заробітна плата млн грн; x_{06} – індекс споживчих цін; x_{07} – видатки бюджету, млн грн; x_{08} – податкові надходження; x_{09} – обсяг депозитів в економіці; x_{10} – облікова ставка НБУ; x_{11} – середньозважена ставка за всіма інструментами; x_{12} – експорт, млн дол; x_{13} – імпорт, млн дол; x_{14} – офіційний курс гривні до долара (середній за період), грн; x_{15} – офіційний курс гривні до євро (середній за період), грн; x_{16} – відсоткова ставка за депозитами, %; x_{17} – обсяг наданих банками кредитів, млн грн; x_{18} – введено в дію нових основних засобів, млн грн; x_{19} – знос основних засобів, %; x_{20} – вибуло основних засобів, млн грн; x_{21} – амортизація, млн грн; x_{22} – вартість основних засобів, млн грн; x_{23} – капітальні інвестиції, всього, млн грн; x_{24} – капітальні інвестиції в харчову промисловість, млн грн; x_{25} – інвестиції в основний капітал сільського господарства, млн грн; x_{26} – валова додана вартість сільського господарства у фактичних цінах, млн грн; Y – валовий регіональний продукт у фактичних цінах, млн грн. Топологія побудованої мережі Байєса представлена на рис. 6.2

Для побудови топології мережі Байєса, розробленої для сценарію S3 використано програмне забезпечення GeNIe 2.0 з такими налаштуваннями:

- метод побудови K2;
- максимальна кількість предків – 5;
- критерій оптимізації – мінімальна ентропія.

Оцінка параметрів моделі виконана на основі рекурсивного методу найменших квадратів

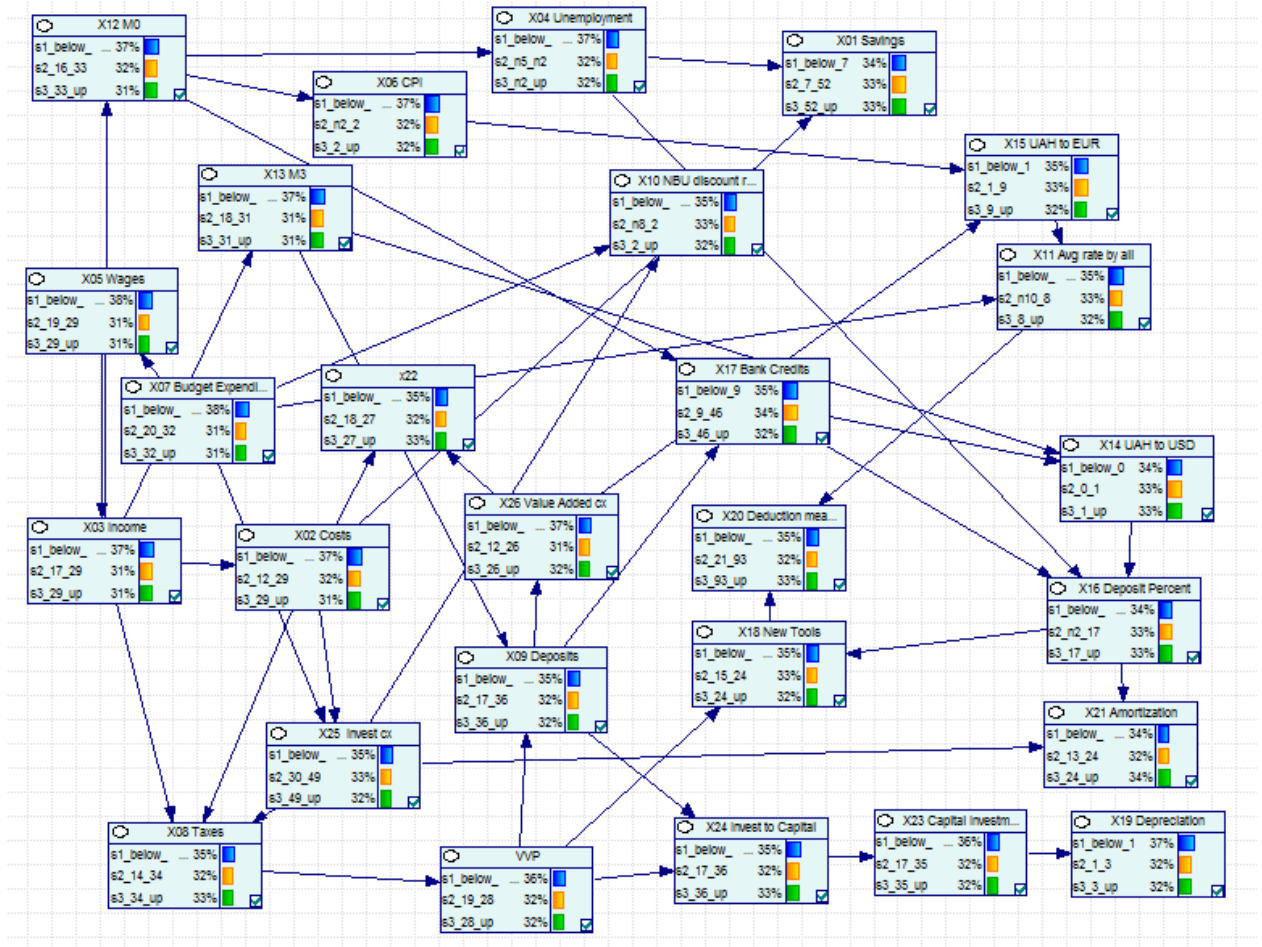


Рис. 6.2. Топологія мережі Байєса для сценарію S3

Таблиця 6.4. Значення статистичних характеристик тестової вибірки наївної мережі Байєса та за допомогою методу K2 в системі GeNIe 2.0.

Назва моделі	Відсоток неправильно класифікованих спостережень	ROC індекс	Коефіцієнт GINI
Наївна мережа	21,3	0,72	0,44
K2 GeNIe 2.0	12,5	0,81	0,62

Як можна побачити з отриманих результатів моделювання, наведених в таблиці 6.4, більш складний підхід із використанням спеціалізованого програмного забезпечення GeNIe 2.0, на основі методу K2, у порівнянні із найпростішим випадком використання мереж Байєса у вигляді наївної мережі, дає більш якісні прогнозуючі характеристики. В якості тестової вибірки для

навчання та тестування використовувалися статистичні дані за період 2003 – 2013 роки, що описуються економічні та соціальні процеси в сільському господарстві.

На основі побудованої топології виявлено, що за реалізації сценарію S3 найбільшого приросту валового регіонального продукту (2 %) можна досягти за умови зростання таких чинників, як податкові надходження (більше ніж на 30 %), обсяг депозитів в економіці (більше ніж на 20 %), капітальних інвестицій в сільське господарство (більше, ніж на 36 %), зростання обсягів валової доданої вартості продукції сільського господарства (більше ніж на 25 %), зростання обсягів капітальних інвестицій в харчову промисловість (більше ніж на 20 %). Ймовірність реалізації такого сценарію досить низька – 26 %. Якщо ж зростуть податкові надходження (більше ніж на 14 %), обсяг депозитів в економіці (не менше ніж на 17 %) та інвестицій в основний капітал (більше, ніж на 36 %), валова додана вартість продукції сільського господарства (більше ніж на 15 %), обсяги капітальних інвестицій в харчову промисловість (більше ніж на 10 %), то валовий внутрішній продукт зросте на 2,8 %. Ймовірність реалізації такого розвитку подій – 51 %.

6.2 Застосування інформаційних технологій прогнозування нелінійних процесів в енергетиці

Впровадження світової практики прогнозного моделювання енергоринку на середньо- та довгострокову перспективу потребує адаптації до національних умов. Тому, враховуючи, що процеси в енергетичній галузі є нелінійними, характеризуються нестационарністю, для прогнозування енергетичною складової запропоновано такі інструменти аналізу, як аналіз часових рядів, математичного моделювання, графічний аналіз.

На першому етапі, використовуючи графічний та статистичний аналіз досліджено наявність типових патернів та відмінностей поведінки системи за різних умов. Було з'ясовано, що найменше навантаження на енергосистему - у 4-5 годинах ночі, а найбільше з 11 до 20 години (рис. 6.3).

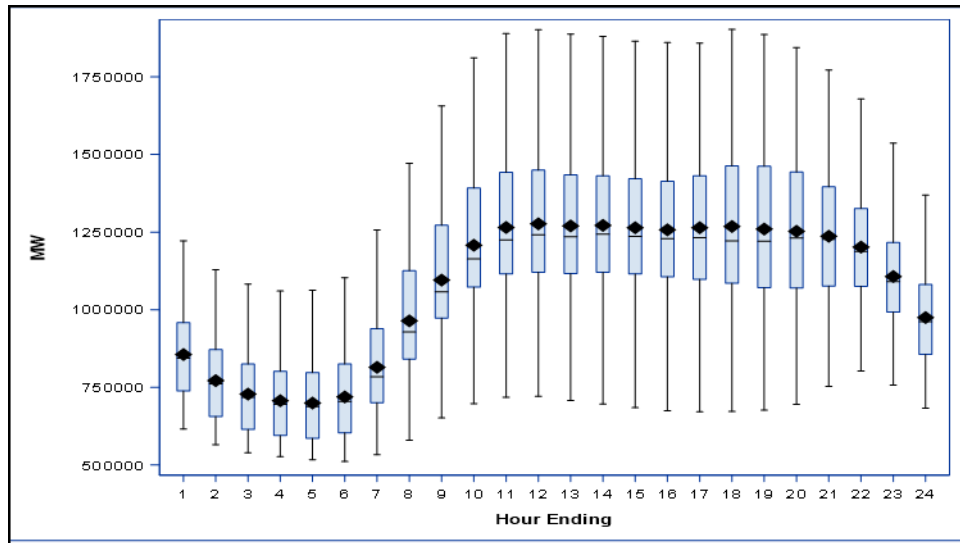


Рисунок 6.3 – Вох-графік погодинного навантаження енергосистеми (ось абсцис – години доби, ось ординат – навантаження енергосистеми, МВт)

Для підвищення якості прогнозу у відповідних математичних моделях для прогнозування навантаження енергосистеми враховано цикли: годинні, добові, місячні та річні, а також, зміни енергоспоживання населенням та бізнесом у передсвяткові, святкові та післясвяткові дні, час доби, температуру навколишнього середовища, географічне розташування споживачів.

Важливим питанням є й прогнозування пікового навантаження, адже за відповідної цінової політики, можна мотивувати споживачів скорочувати споживання електроенергії у пікові години або, вчасно реагувати на загрози енергосистемі. У даному прикладі розглянуто часові ряди навантаження енергосистеми, побудовано є авторегресійні моделі.

Модель дванадцятого порядку для випадку прогнозування місячних даних в енергетиці має наступний вигляд:

$$y(t) = 602546044 - 0,0219 \cdot y(t-1) - 0,0257 \cdot y(t-2) - 0,0548 \cdot y(t-3) - 0,1528 \cdot y(t-4) \\ - 0,0029 \cdot y(t-5) - 0,0886 \cdot y(t-6) - 0,0184 \cdot y(t-7) - 0,1687 \cdot y(t-8) - 0,0642 \cdot y(t-9) + \\ + 0,0013 \cdot y(t-10) + 0,0827 \cdot y(t-11) + 0,7646 \cdot y(t-12)$$

де $y(t)$ - навантаження енергосистеми за місяць.

Моделі класу експоненційного згладжування також широко використовуються для моделювання часових рядів, а для випадку наявності тренду та сезонності опрацьовані їх відповідні модифікації. Загальний вигляд моделі (6.4):

$$\hat{y}(t+h) = (a(t) + h \cdot b(t)) \cdot s(t-p+1+(h-1)\text{mod}(p)) \quad (6.4)$$

де $\hat{y}(t+h)$ – прогноз на період h вперед, p – довжина періоду сезонності, mod – функція визначення залишку від ділення, $a(t)$ ($a(t) = \alpha \cdot \left(\frac{y(t)}{s(t-p)}\right) + (1-\alpha) \cdot (a(t-1) + b(t-1))$) – рівняння, що описує згладжений тренд, рівняння $b(t)$ ($b(t) = \beta \cdot (a(t) - a(t-1)) + (1-\beta) \cdot b(t-1)$), що використовується для оцінювання тренду, $s(t)$ - ($s(t) = \gamma \cdot \left(\frac{y(t)}{a(t)}\right) + (1-\gamma) \cdot s(t-p)$) рівняння для оцінювання сезонності.

Модель для прогнозування щоквартального навантаження енергосистеми має вигляд:

$$y(t) = 1\,861\,428\,237 + 0,03931 \cdot y(t-1) - 0,4818 \cdot y(t-2) + 0,0833 \cdot y(t-3) + 0,5842 \cdot y(t-4)$$

де $y(t)$ – навантаження енергосистеми за квартал.

Серед методів класу експоненційного згладжування найкращі результати прогнозування отримані за моделлю Вінтерса з мультиплікативною сезонною складовою. В таблиці 6.5 наведені параметри згладжування для моделі Вінтерса з мультиплікативною сезонністю.

Таблиця 6.5. – Значення параметрів згладжування для моделі Вінтерса з мультиплікативною сезонністю

Параметр	Оцінка	Стандартна похибка	Статистика Стьюдента	<i>p-value</i>
α	0,4294	0,1272	3,3767	0,0019
β	0,0983	0,0327	3,0077	0,0050
γ	0,1298	0,0262	4,9467	0,0000

В таблиці 6.6 представлені отримані значення оцінок коефіцієнтів рівняння з відповідними статистичними характеристиками.

Таблиця 6.6. – Порівняння статистичних характеристик авторегресійної моделі AP(12) та моделі Вінтерса

Назва показника	AP(12)	AP(4)	Модель Вінтерса з мультиплікативною сезонністю
Коефіцієнт детермінації (R^2)	0,95	0,98	0,83
Середньоквадратична похибка (RMSE, млн МВт)	24,866	39,504	24,873
Середня абсолютна процентна похибка (MAPE)	2,64%	6,9 %	2,6%

Отже, при середньостроковому прогнозуванні, на місячних відрізках часових рядів кращі результати одержано за моделлю авторегресії дванадцятого порядку та моделлю Вінтерса з мультиплікативною сезонністю, для яких значення статистики MAPE становить 2,64% та 2,6% відповідно. Менш точні результати було отримано при короткостроковому прогнозуванні на квартальних даних, похибка авторегресії четвертого порядку сягнула значення статистики MAPE в 6,9%. Аналіз даних в добовому та місячному розрізах «температура-навантаження» засвідчив залежність у вигляді оберненої фази в залежності від доби року, тому, значення температури повітря необхідно враховувати в моделі у вигляді поліному третього порядку разом із спеціалізованими комбінованими змінними на основі температури та місяця, температури та години доби, температури та дня тижня. Більш детальний аналіз результатів показав, що загалом 93% попереднього значення навантаження враховується в наступному, при цьому присутній тренд в часових даних, а підмішування в модель регресорів що описують свята підвищує точність прогнозування за критерієм MAPE на 0,04%.

6.3 Прогнозування еколого-економічних складових регіонального розвитку

Для більшості регіонів України, аграрна галузь є провідною ланкою їх економіки, тому питання прогнозування урожайності польових

сільськогосподарських культур у певній агрокліматичній зоні за комбінованого впливу біологічних, екологічних, економічних та антропогенних складових є актуальним, має практичне значення як з точки зору наповнення місцевого бюджету, та і забезпечення продовольчої безпеки. Враховуючи результати проведених досліджень, слід відзначити, що найкращі результати можна одержати опрацьовуючи набори даних урожайності польових культур значної кількості господарств певної агрокліматичної зони та застосовуючи композиції різних методів прогнозного моделювання, уточнюючи при цьому побудовані моделі відповідно до реалій протікання процесу, що досліджується. Тобто у системі підтримки прийняття рішень сільськогосподарського призначення для прогнозування потенціалу урожаю необхідно застосовувати методику адаптивного моделювання із попередньою підготовкою даних.

Застосування такої методики значно розширює можливості систем підтримки прийняття рішень сільськогосподарського призначення за рахунок реалізації широкого спектру моделей, серед яких такі: лінійні та нелінійні регресійні моделі (зокрема, $AR(p)$, $ARCS(p,q)$, $HARCS$, $ARUG$, $UARUG$, $E-UARUG$); поліноміальні, експоненціальні та інші види трендів; логістична регресія; динамічні і статистичні байєсівські мережі, дерева рішень та деякі інші. Попередня підготовка даних до побудови математичної або статистичної моделі включає у себе такі обчислювальні операції: нормування у вибраному діапазоні, заповнення пропусків з використанням альтернативних методів, цифрову та оптимальну фільтрацію, розмноження вибірки, тестування на стаціонарність, наявність коінтегрованості і т. ін. Для аналізу якості даних використовується множина статистичних критеріїв якості, які необхідні для перевірки інформативності вибірки, наявності екстремальних значень, оцінювання рівня шумових компонент та визначення кількості пропусків. Отримані результати стосовно якості даних використовуються для покращення їх характеристик.

Наступний етап після встановлення класу процесу (лінійний стаціонарний або нелінійний нестаціонарний) стосується оцінювання структури і параметрів моделі у вибраному класі. Структури моделей-кандидатів (порядок та

розмірність моделі, наявність нелінійності, оцінка лагу, тип розподілу збурення та його параметри) оцінюються на основі аналізу статистичних характеристик процесу, законів та закономірностей його протікання, а також накопичених знань стосовно аналізу процесів конкретного класу. Після оцінювання структури оцінюються параметри моделей-кандидатів. Лінійні та псевдолінійні моделі оцінюються за звичайним методом найменших квадратів (МНК), а моделі, нелінійні стосовно параметрів, оцінюються за нелінійним МНК, методом максимальної правдоподібності (ММП), коли це можливо, або за методом Монте-Карло для марковських ланцюгів (МКМЛ). Останній метод – універсальний для різних типів розподілів даних і типів нелінійності. Прикладом запропонованої реалізації даного адаптивного системного підходу є дослідження можливості отримання приросту урожайності озимої пшениці залежно від агрометеорологічних умов. Вхідними даними є показники урожайності озимої пшениці в зоні північного Лісостепу України та метеорологічні показники Київської метеостанції. Попередньо вхідні дані перевірені на наявність шумів, пропусків, екстремальних значень, тощо. Враховуючи особливості процесу, що досліджується, зокрема наявність значної кількості факторів, що впливають на зміну урожайності польової культури, для прогнозування урожайності було побудовано ряд моделей, що відрізняються за статистичними критеріями. В якості середовища моделювання використано SAS Enterprise Miner. Зокрема, для визначення взаємовпливу факторів, виконано двофакторний дисперсійний аналіз з перетином змінних, в результаті якого виявлено, що для підвищення якості прогнозування при побудові моделі крім попередника та обсягу внесених добрив слід враховувати температуру та кількість опадів у окремі періоди вегетації рослин. Оскільки сприятливі значення агрокліматичних факторів і тривалість фаз розвитку рослин відомі за багаторічними спостереженнями, а сівозміна передбачена агротехнологією вирощування культури, то для прогнозування урожайності озимої пшениці побудовані відповідні регресійні моделі. Статистичні характеристики побудованих регресійних моделей подані у табл. 6.7.

Таблиця 6.7 – Статистичні характеристики регресійних моделей, побудованих для прогнозування урожайності озимої пшениці

Фактори моделі множинної регресії	RMSE	MAPE	DW
Тільки добрива	1,19	24,78	2,25
Тільки культура попередник	1,24	26,13	2,07
Добрива та культура попередник	1,17	24,59	2,31
Добрива, культура-попередник, добрива-культура-попередник-ріст-регулятор	1,12	24,59	2,32
Тільки температура	0,93	20,58	1,41
Тільки опади	1,05	19,43	1,48
Тільки температура та опади	0,69	13,62	1,17
Добрива, культура попередник, температура та опади	0,52	10,01	1,92

Для коректного включення в модель регресорів температура та опади запропонована методика перетворення регресорів, що реалізується за кілька кроків.

Крок 1. Обчислюються середні значення температури та опадів у відповідні місяці вегетації.

Крок 2. На основі отриманих середніх значень обчислюються відхилення реального значення температури: $\Delta T_i = T_i - \underline{T}$, де ΔT_i – відхилення температури T_i у i -му місяці від значення математичного сподівання $\underline{T}$, що обчислене на основі даних з 1988 по 2020 роки. та опадів $\Delta P_i = P_i - \underline{P}$, де ΔP_i – відхилення кількості опадів P_i у i -му місяці від значення математичного сподівання $\underline{P}$, що обчислене на основі даних з 1988 по 2020 рр. у відповідному році для вказаної агрокліматичної зони.

Вибір регресорів відбувався в автоматичному режимі за методом Stepwise з пороговими значеннями, рівними 0,15 входу та збереження регресора у моделі. В табл. 6.8 подані значення оцінок параметрів лінійної регресії, що увійшли в модель.

Таблиця 6.8 - Значення оцінок окремих параметрів лінійної регресії, що увійшли в модель

Змінна регресор / Параметр	DF	Оцінка	Standard Error	WaldChi-Square	Pr > ChiSq
Вільний член	1	3,3880	0,0761	1984,26	<,0001
T ₁₀	1	-0,6320	0,0482	171,78	<,0001
T ₁₂	1	0,3202	0,0149	458,70	<,0001
T ₂	1	0,2075	0,0198	109,26	<,0001
T ₃	1	-0,2754	0,0222	153,38	<,0001
P ₉	1	0,0099	0,0008	139,49	<,0001
P ₃	1	-0,0219	0,0018	145,19	<,0001
Добриво N60P40K40	1	0,9216	0,0909	102,77	<,0001
Гній (30 т/га)	1	0,9914	0,0909	118,94	<,0001
Добриво (гній 30т/га+N60P40K40)	1	1,1130	0,0909	149,89	<,0001
Добриво = N120P80K80	1	1,0637	0,0909	136,92	<,0001
Добриво = контрольна	0	0,0000	0,0000	.	.
Культура-попередник горох	1	0,4456	0,0704	40,04	<,0001
Культура-попередник багаторічні трави	1	0,2983	0,0704	17,95	<,0001

Виходячи з отриманих в ході побудови моделей результатів, слід відзначити, що досягти показника потенційної урожайності сорту для озимої пшениці можна за різних сценаріїв використання добрив, попередників, метеорологічних факторів. Наприклад, як показали розрахунки: використання добрива N60P40K40 збільшує урожайність культури на 0,9216 ніж в контрольному випадку без добрива, добрива N120P80K80 – на 1,0637 ніж в контрольному випадку, попередник горох, сприяє підвищенню урожайності на 0,4456, на відміну від попередника кукурудза на силос, збільшення температури в жовтні на 1 градус від середнього значення 8,53 зменшує продуктивність на 0,632, збільшення опадів у вересні на 1 одиницю від середнього значення 62,82 збільшує продуктивність на 0,0099, тощо.

Тому з метою оцінювання ймовірності можливості зростання урожайності озимої пшениці під впливом різних комбінацій культур у сівозміні та метеорологічних факторів, використано ймовірнісне моделювання. Переваги байєсівського підходу полягають у тому, що апріорні ймовірності можна уточнювати відповідно до змін у протіканні процесу, що досліджується. Це надає

можливість уточнювати ймовірності подій при надходженні додаткової інформації і створювати системи адаптивного моделювання. При побудові структури мережі Байєса опрацьовано таку множину станів: стан L2 (якщо значення відхилення менше або рівне значення одного стандартного відхилення); стан L1 (якщо значення відхилення більше одного стандартного відхилення та менше нуля); стан AV (якщо значення відхилення близько до нуля); стан H1 (якщо значення відхилення більше нуля та при цьому менше значення одного стандартного відхилення); стан H2 (якщо значення відхилення більше або дорівнює значенню одного стандартного відхилення).

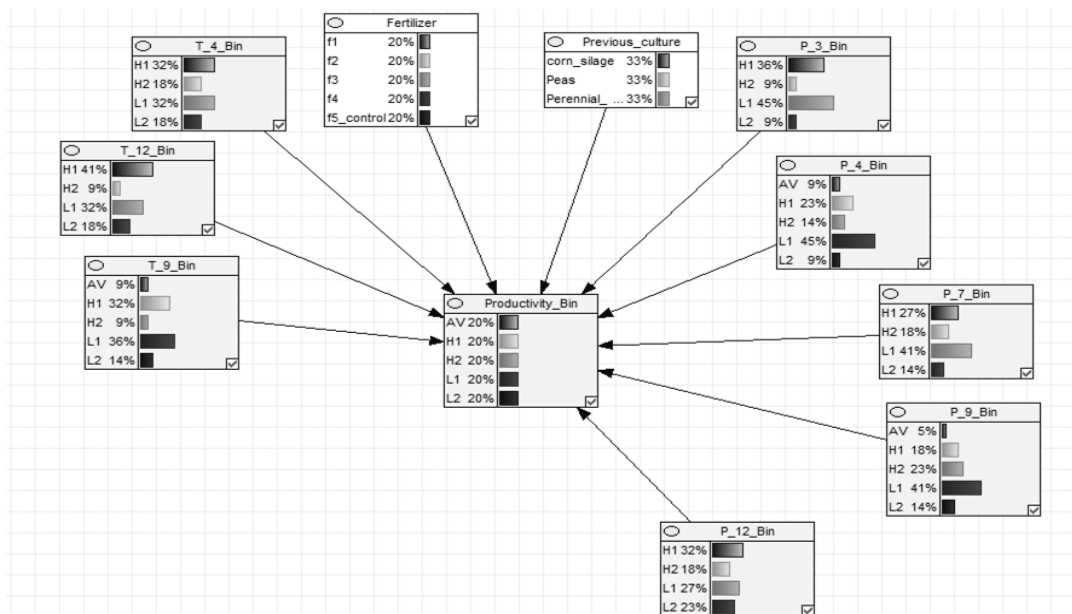


Рисунок 6.4 – Структура мережі Байєса для прогнозування зміни урожайності озимої пшениці

Сценарій 1. Для забезпечення максимальної урожайності вершини мережі Байєса повинні мати такі значення. Температура у квітні (значення H1) повинна бути вищою середньої (більше 9,5 градусів та менше 11,1, у вересні (значення H1) – вище середньої (тобто 14,3 градусів та менше 15,9, у грудні (значення L1) – нижче середнього (тобто більше -5,4 градусів та менше -2,6 тому, що $\text{MX} = -2,5$, $\text{STD} = 2,8$); опади у березні (значення L1) повинні бути нижче середнього значення (тобто більше 13 мм та менше 33мм, у квітні (значення L1) – нижче середнього (тобто більше 13 мм та менше 39 мм; тому, що $\text{MX}=40$, $\text{STD}=26$), у липні (значення L1) – нижче середнього (більше 35 та менше 70 мм; тому що $\text{MX}=71$, $\text{STD}=36$); у вересні (значення L1) – нижче середнього (тобто більше 17

та менше 62; тому що $MX=63$, $STD=46$); грудень (значення $H1$) – вище середнього (тобто більше 17 та менше 37; тому що $MX=38$, $STD=21$).

Сценарій 2. Якщо вершини перевести у такі стани: «Температура»: квітень, вересень, грудень (значення $H2$) значно вище середнього; «Опади»: березень, квітень, липень, вересень, грудень (значення $H2$) стан: значно вище середнього; «Добрива» – контрольне значення (без добрива), культура попередник – кукурудза на силос, то з ймовірністю 31% буде отримане максимальний приріст врожайності – $H2$.

Сценарій 3. Якщо вершини перевести у стани: «Температура»: квітень (значення $H1$) – вище середнього, вересень (значення $H2$) значно вище середнього, грудень (значення AV) – середнє значення (мінус 2,5 градусів Цельсія), «Опади»: березень (значення $L2$), квітень (значення $L2$), липень (значення $L2$), грудень (значення $L2$) значно нижче середнього, а вересень (значення AV) – середнє значення (середнє значення 63), «Добрива» (значення $F4$) – N120P80K80, «Культура попередник» – багаторічні трави, тоді з ймовірністю у 45% буде отримано максимальний приріст врожайності ($H2$).

Наведений приклад застосування адаптивного підходу до прогнозного моделювання урожайності сільськогосподарських культур свідчить про хороші перспективи його практичного використання для прогнозування урожайності польових сільськогосподарських культур з метою забезпечення раціонального природокористування.

6.4 Інформаційна технологія класифікації об'єктів для цілей регіонального розвитку

Підхід щодо обчислення міри подібності часових рядів, може бути застосований і в складі ансамблів моделей. Зокрема, у випадку, коли дані неповні або є сумніви щодо їх достовірності, доцільно застосовувати методику, що має в основі пошук історичних аналогій поведінки досліджуваних процесів. В цьому випадку необхідна історична вибірка даних та відрізок-шаблон для якого виконується пошук історичного аналогу. Зазвичай для таких задач виміри

прив'язані до часу. Реалізація пропонованої методики виконується після попередньої обробки даних із виконанням описаних у попередніх розділах процедур. Подальші дії виконуються у наступній послідовності:

Крок 1. Робота з вхідними даними: часовим рядом історичних даних та часовим рядом шаблону, який буде виконувати роль цільового ряду, і з яким буде виконуватися порівняння часового ряду історичних даних.

Крок 2. Визначення часового інтервалу на якому буде виконуватися аналіз та агрегування даних. (Виконується для випадку, коли дані мають вигляд транзакцій. Тоді їх необхідно агрегувати до заданого часового інтервалу, наприклад, секундні виміри привести до хвилинних агрегованих значень, або до погодинних і так далі).

Крок 3. Ідентифікація або задання користувачем (аналітиком) значення тривалості сезонного циклу (більшість часових рядів, що описують реальні процеси характеризуються наявністю сезонних складових).

Крок 4. Розбиття вхідного часового історичний ряду даних на відрізки відповідної довжини. В результаті буде отримано K вхідних історичних часових рядів.

Крок 5. Обчислення значень міри подібності між кожним з вхідних часових рядів, отриманих на попередньому кроці (крок 4), та цільовим часовим рядом. Результати будуть мати вигляд колектору значень подібностей:

$$Sim = \{Sim_1, \dots, Sim_k\}$$

де Sim_i -- міра подібності між i -м вхідним інтервалом та цільовим рядом, що обчислюється за формулою (6.5):

$$ScoreSim(Y, X) = 100\% \cdot \left(1 - \frac{ScoreAvg_1}{ScoreAvg_0}\right) == 100\% - 100\% \times \frac{\left(\frac{1}{(max(n,m)-(P_1-max(n,m)))} \sum_{k=1}^{P_1} D_1[j_p, i_p]\right)}{\left(\frac{1}{n} \sum_{k=1}^{P_0} D_0[j, j]\right)} \quad (6.5)$$

Крок 6. Пошук оптимального значення подібності, серед усіх значень вектору отриманого на попередньому кроці (6.5):

$$\max \{Sim_i : i = 1, \dots, K\}$$

Слід зазначити, що у загальному випадку, замість міри подібності, формула (5.1), може бути використана будь-яка інша метрика – наприклад стандартна ScoreAvg, тоді відповідно на кроці 6 вирішується задача мінімізації – тобто визначається пара рядів для яких значення ScoreAvg найменше:

$$\min \{ScoreAvg_i : i = 1, \dots, K\}$$

$$ScoreSim(Y, X) = 100\% \cdot \left(1 - \frac{ScoreAvg_1}{ScoreAvg_0}\right)$$

Після того, як результати сформовані, аналітик на основі побудованих графіків та обчислених статистик, приймає рішення, щодо подальшого використання результатів, або введення додаткової інформації та уточнень в задачу.

Не рідкість, коли на практиці потрібно вирішити задачу класифікації або сегментації об'єктів, осіб, процесів тощо за їхньою поведінкою у часі. Одними із підходів до вирішення цієї проблеми є підхід, оснований на виявленні подібності часових рядів.

У загальному вигляді задачу можна сформулювати наступним чином. Нехай, є K часових рядів, $Y = \{Y^{(1)}, \dots, Y^{(K)}\}$, кожен з яких описує поведінку деякого об'єкта в часі. Тоді, можна побудувати матрицю подібностей поведінки об'єктів, розмірності $K \times K$:

$$Sim = \begin{pmatrix} Sim_{1,1} & \dots & Sim_{1,K} \\ \vdots & \ddots & \vdots \\ Sim_{K,1} & \dots & Sim_{K,K} \end{pmatrix} = \begin{pmatrix} Sim(Y^{(1)}, Y^{(1)}) & \dots & Sim(Y^{(1)}, Y^{(K)}) \\ \vdots & \ddots & \vdots \\ Sim(Y^{(K)}, Y^{(1)}) & \dots & Sim(Y^{(K)}, Y^{(K)}) \end{pmatrix}$$

де $Sim_{i,j} = Sim(Y^{(i)}, Y^{(j)})$ – це значення міри подібності між i – м та j – м часовими рядами, що описують поведінки відповідних об'єктів.

Зауважимо, що матриця має симетричний вигляд, тобто $Sim_{i,j} = Sim_{j,i}$, а діагональні елементи в загальному випадку не обчислюються, тому що ступінь подібності часового ряду з самим собою буде стовідсотковою.

З урахуванням наведених зауважень стає зрозумілим, що потрібно в загальному випадку зробити $\frac{K \cdot (K-1)}{2}$ попарних порівнянь між часовими рядами.

Після того як усі необхідні обчислення виконані та отримана матриця подібності, об'єкти об'єднуються в групи із використанням алгоритму кластеризації. Для визначення кількості кластерів у задачах прогнозування нелінійних нестаціонарних процесів, за допомогою методу мінімальної дисперсії Уорда, k -середніх та інших методів, що базуються на використанні мінімізації суми квадратів усередині кластера для задач, як показали чисельні експерименти доцільно застосовувати критерій кубічної кластеризації (CCC), який розраховується із використанням методів Монте-Карло [188] (6.6):

$$CCC = \ln \left(\frac{1 - E(R^2)}{1 - R^2} \right) \frac{\sqrt{\frac{np^*}{2}}}{(0.001 + E(R^2))^{1.2}} \quad (6.6)$$

Формулу (6.6) було отримано емпірично, намагаючись стабілізувати дисперсію між різною кількістю спостережень, змінних та кластерів [188].

$$E(R^2) = 1 - \left| \frac{\sum_{j=1}^{p^*} \frac{1}{n + u_j} + \sum_{j=p^*+1}^p \frac{u_j^2}{n + u_j}}{\sum_{j=1}^p u_j^2} \right| \cdot \left| \frac{(n - q)^2}{n} \right| \cdot \left| 1 + \frac{4}{n} \right|$$

$$R^2 = 1 - \frac{p^* + \sum_{j=p^*+1}^p u_j^2}{\sum_{j=1}^p u_j^2}$$

$$u^* = \prod_{j=1}^{p^*} s_j$$

$$c = \left(\frac{v^*}{q} \right)^{\frac{1}{p^*}}$$

$$u_j = \frac{s_j}{c}$$

де q – кількість кластерів, з якими порівнюють p^* , v – розмір гіперкуба, c – довжина ребра гіперкуба, який отримують при діленні гіперкуба v на q частин, u_j – кількість гіперкубів вздовж j -го виміру основного гіперкубу, s_j – корінь квадратний з j -го власного числа, що описує множину даних X , n – кількість спостережень.

ССС-критерій використовується для визначені кількості кластерів за методом Уорда, наступним чином:

- послідовно перевіряються гіпотези, про кількість кластерів;
- для кожної гіпотези робиться розбиття на відповідну кількість кластерів та обчислюється значення СССР-критерію;
- фахівець аналізує отримані значення СССР-критерія та приймає рішення щодо відповідної кількості кластерів, які описують набір даних.

Емпірично експерти [188] з'ясували та рекомендують використовувати одне з двох правил.

Правило 1. Обирають ту кількість кластерів, для якої значення СССР-критерія перевищує 3.

Правило 2. Обирають ту кількість кластерів, для якої значення СССР-критерія стрімко зростає, після падіння.

Окрім наведеного вище підходу групування об'єктів на основі методів кластерного аналізу, можуть бути використані й інші методи кластеризації. Наприклад, на основі експертного досвіду аналітика, а саме, після того, як усі необхідні обчислення виконані та отримана матриця подібності, об'єкти об'єднуються в групи, за правилом: якщо міра подібності $Sim_{i,j} = Sim(Y^{(i)}, Y^{(j)})$ більша за деяке порогове значення, то i -й та j -й об'єкти, що описуються відповідно часовими рядами $Y^{(i)}$ та $Y^{(j)}$ належать одній групі.

Застосування пропонованого методу кластеризації в рамках багатомодельного підходу доцільно продемонструвати на прикладі вирішення

практичної задачі групування регіонів України обсягом капітальних інвестицій, спрямовуваних на охорону навколишнього середовища.

Вхідними даними є статистичні показники обсягу капітальних інвестицій в охорону навколишнього природного середовища в регіонах України за 2011-2018 рр. [173]. Всього досліджується 26 часових рядів. Для обчислення значень схожості рядів виконувалася попередня нормалізація даних, а саме, нормування за наступною формулою:

$$z_i = \frac{x_i - \min_{i=1,..,m} \{x_i\}}{\max_{i=1,..,m} \{x_i\} - \min_{i=1,..,m} \{x_i\}}$$

В якості порогового значення схожості рядів за метрикою *ScoreSim* обрано значення 70%, тобто на графі зав'язків відображаються тільки ті зв'язки між областями що дорівнюють або перевищують 70% (найбільш суттєві).

На основі значень схожості рядів, отриманих за метрикою *ScoreSim*., використовуючи будь-який з методів кластеризації, можна згрупувати регіони, за ступенем подібності (близькості за метрикою *ScoreSim*).

Таблиця 6.7 – Результати кластерного аналізу, виконаного за методом Уорда (Ward) регіонів

Регіон	Номер кластеру	Кількість регіонів
Дніпропетровська, Запорізька, Київська, Миколаївська, Сумська, Тернопільська, Хмельницька, Черкаська області, місто Київ	1	9
Волинська, Івано-Франківська, Херсонська, Чернігівська, Чернівецька області	2	5
Кіровоградська, Луганська, Одеська, Рівненська, Запорізька області	3	5
Вінницька, Донецька, Закарпатська, Львівська, Полтавська, Харківська області	4	6

Код програми для імпортування даних та побудови дендограми за методом Уорда (Ward), в системі SAS Studio.

```
proc import datafile=...\Similarities.xlsx'
out=Similarities
dbms=XLSX
replace;
sheet='Similarity';
run;

ods graphics on;

proc cluster data=work.Similarities outtree=tree
method=ward plots=dendrogram;
id _input_;
run;
```

Графічно результати кластеризації представлені на рис. 6.5.

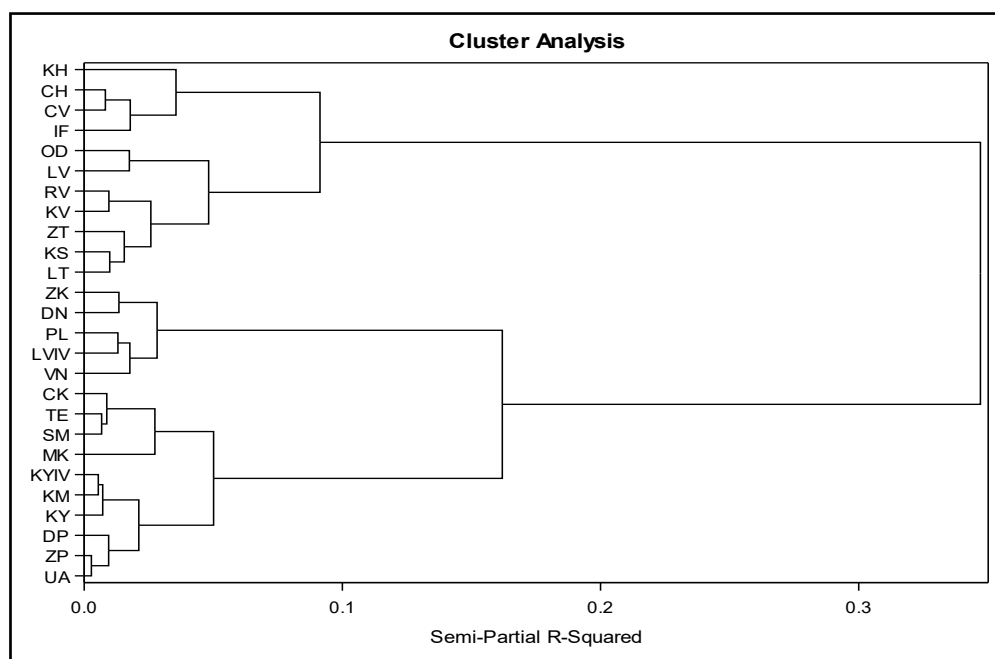


Рисунок 6.5 – Дендограма групування регіонів за обсягом капітальних інвестицій в охорону навколишнього середовища у чотири кластери методом Уорда (Ward), виконана в системі SAS Studio

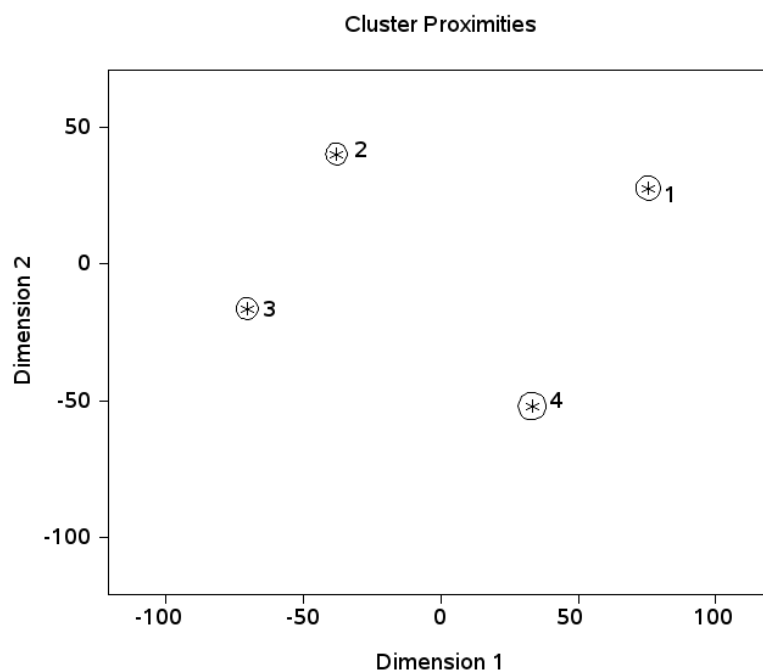


Рисунок 6.6 - Відстань між кластерами, результати роботи системи SAS Enterprise Miner

Для візуалізації результатів кластеризації було використано програму SAS Graph Network Visualization WorkShop (рис. 6.7).

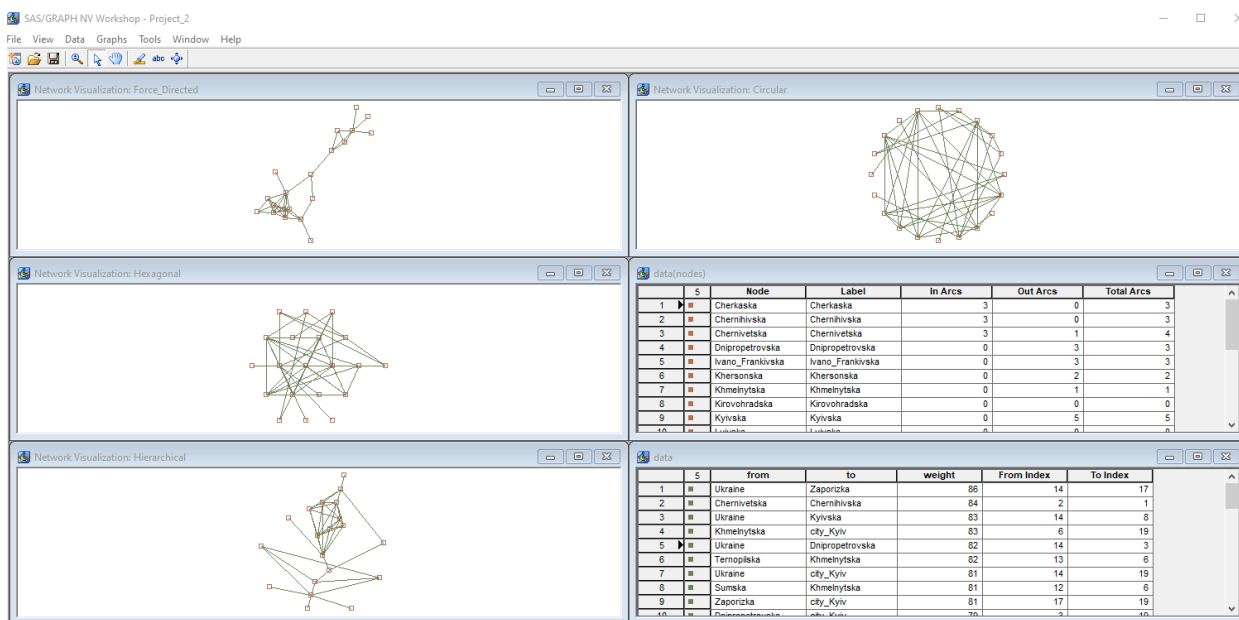


Рисунок 6.7 – Скріншот робочого вікна системи SAS Graph Network Visualization WorkShop

Отримано представлення у вигляді кругового графу; ієрархічного графу; гексагонального графу та багаторівневого графу.

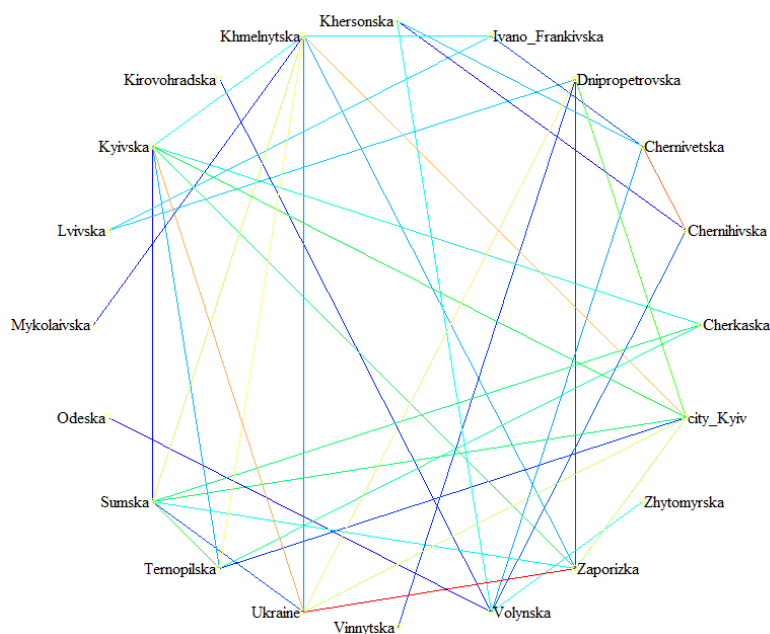


Рисунок 6.8 - Круговий граф зв'язків відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%. Колір залежить від значення метрики ScoreSim

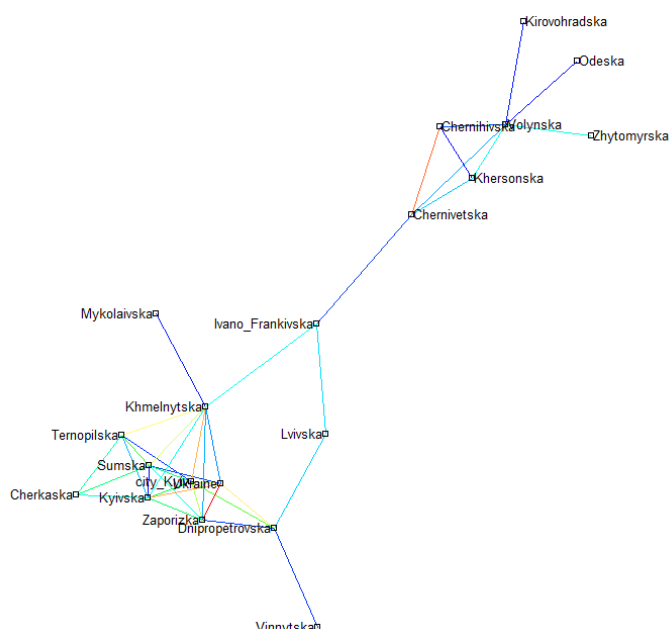


Рисунок 6.9 – Граф зав'язків за координатами користувача, відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%. Колір залежить від значення метрики ScoreSim

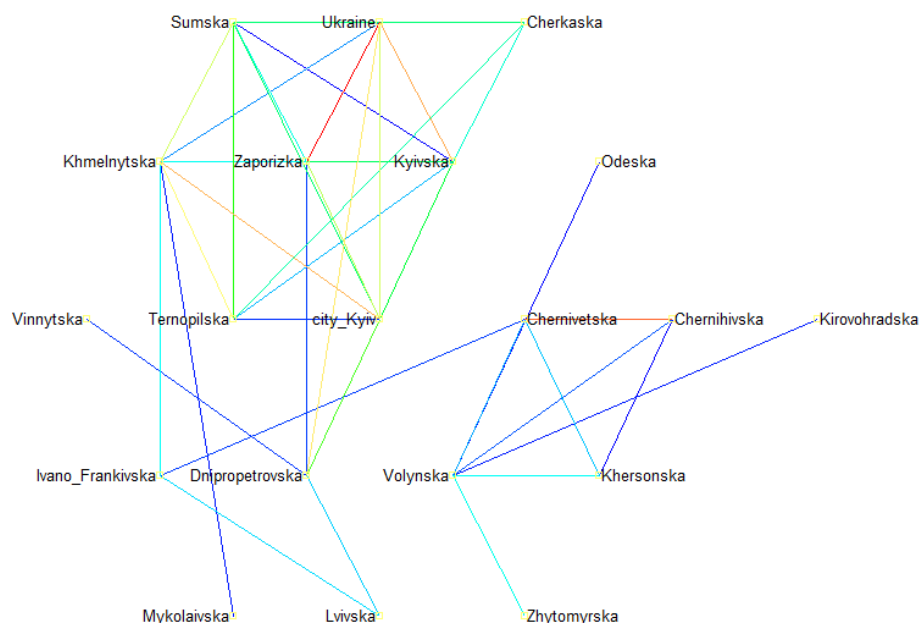


Рисунок 6.10 – Гексагональний граф зв'язків відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%. Колір залежить від значення метрики ScoreSim

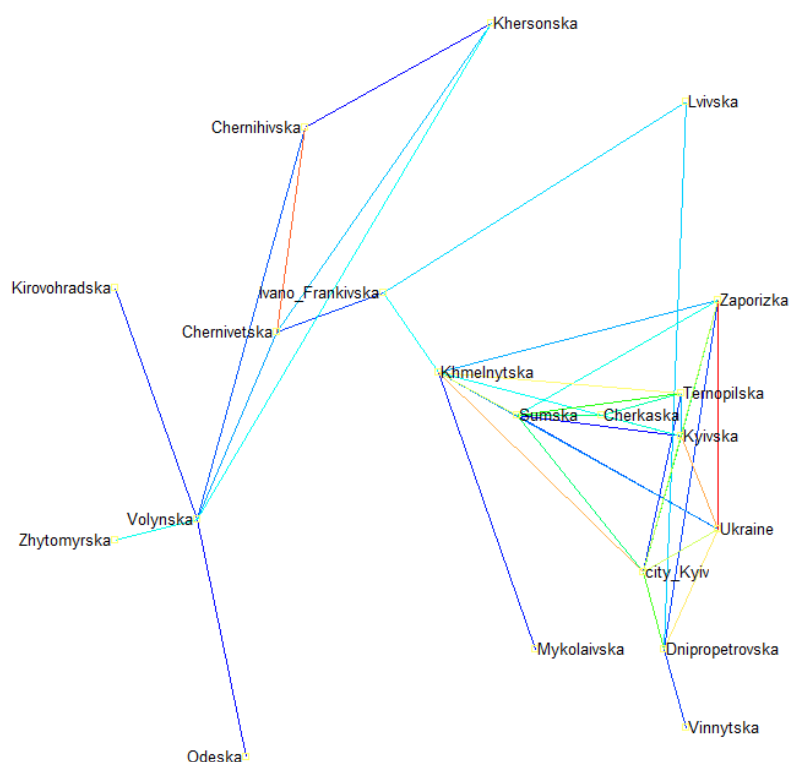


Рисунок 6.11 – Ієрархічний граф зв'язків відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%. Колір залежить від значення метрики ScoreSim

Рис. 6.7 – 6.11 відображають одну і ту саме мережу, але з різним розташуванням вершин, при чому колір залежить від значення метрики ScoreSim.

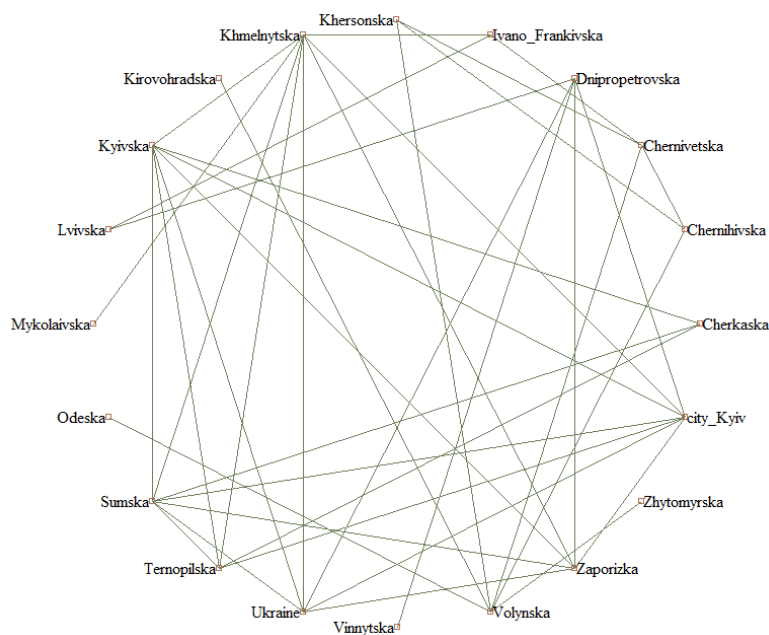
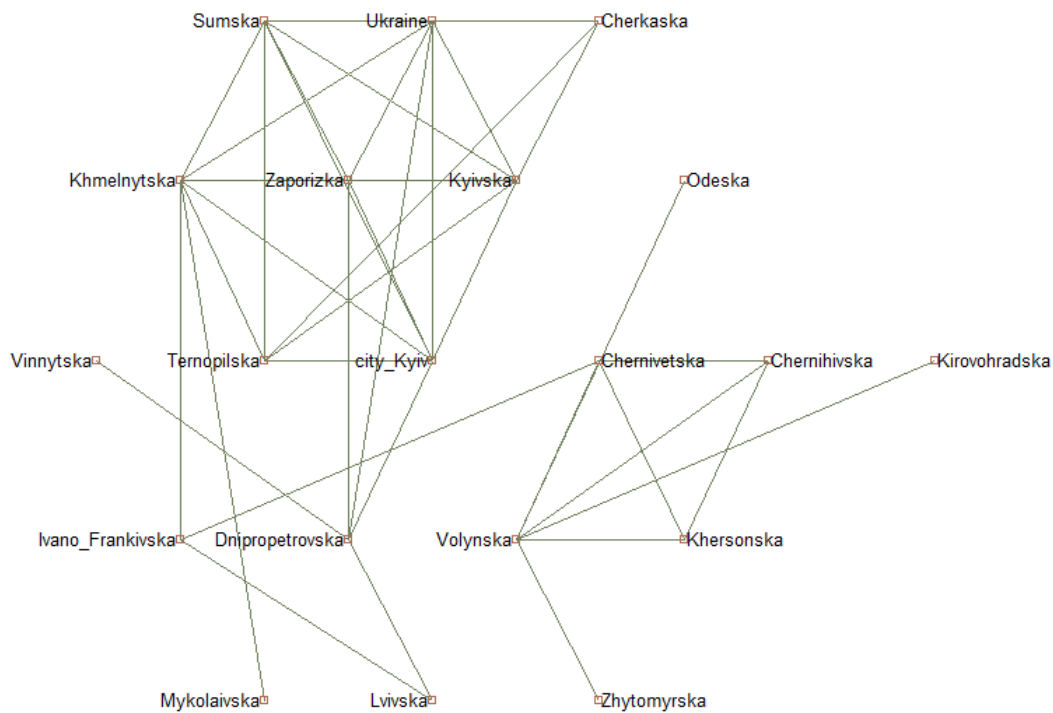


Рисунок 6.12 – Круговий граф зв'язків відображає тільки ті зв'язки між



областями що дорівнюють або перевищують 70%

Рисунок 6.13.- Граф зв'язків, за координатами користувача, відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%

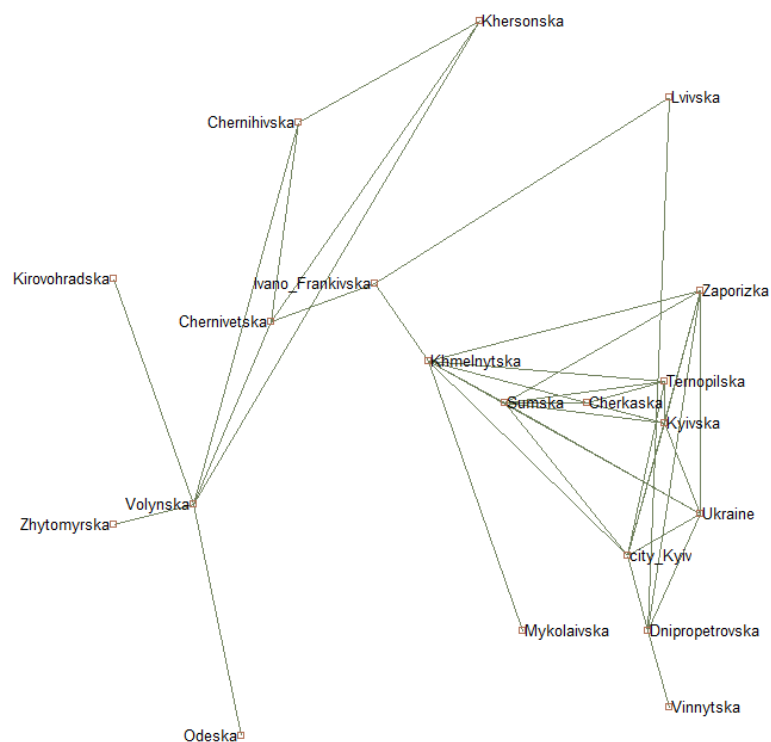


Рисунок 6.14 – Гексагональний граф зв'язків відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%

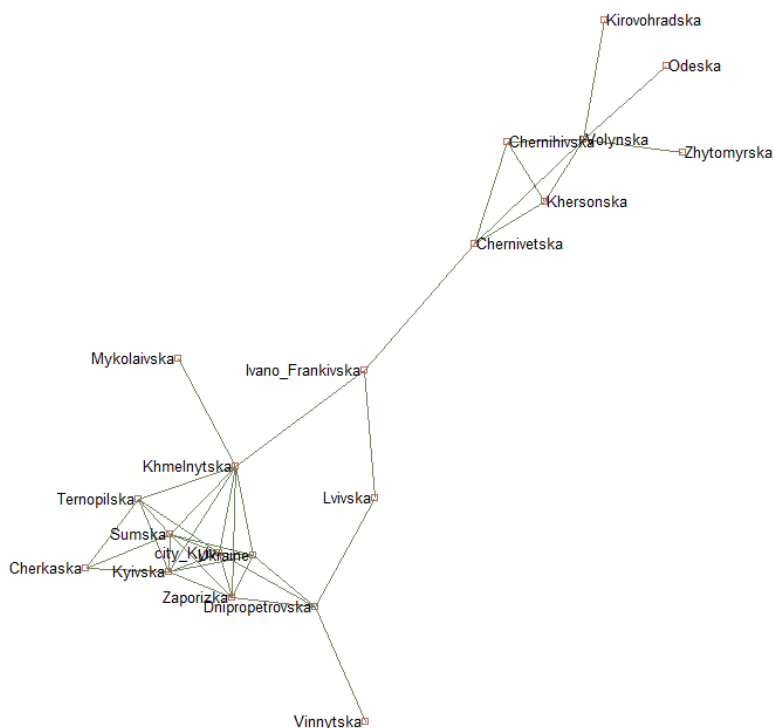


Рисунок 6.15 – Ієрархічний граф зв'язків відображає тільки ті зв'язки між областями що дорівнюють або перевищують 70%

Як можна побачити з отриманих результатів, між значеннями ступеня зв'язку для рядів за кореляцією Пірсона та метрикою подібності рядів ScoreSim є відмінності, насамперед для випадків, коли значення кореляції хоча і великі, але від'ємні, метрика ScoreSim прямує до нуля.

На основі результатів кластеризації, щодо приналежності до кластеру регіону, за результатами аналізу методом Уорда (Ward) в системі SAS Enterprise Miner, можна зробити висновок, що:

- Дніпропетровська, Запорізька, Київська, Миколаївська, Сумська, Тернопільська, Хмельницька, Черкаська області та місто Київ належать до першого кластеру;
- Волинська, Івано-Франківська, Херсонська, Чернівецька та Чернігівська належать до другого кластеру;
- Житомирська, Кіровоградська, Луганська, Одеська та Рівненська належать до третього кластеру;
- Вінницька, Донецька, Закарпатська, Львівська, Полтавська та Харківська до четвертого кластеру.

Отже, використовуючи методику кластеризації даних, оснований на теорії подібних процесів, отримано результати, які можна використовувати в ансамблях моделей, в тому числі і в якості змінної, що включена в модель як змінна, запропонована користувачем.

Урахування особливостей функціонування об'єктів критичної інфраструктури, їх геопросторового розташування, їх технічних, технологічних, особливостей, тощо, під час паспортизації об'єктів, потребує впровадження відповідних підходів до їх класифікації.

Кожний секторальний орган розробляє свої рекомендації, інструкції, нормативи щодо безпекових, організаційних, технічних, технологічних, фінансово-економічних заходів для об'єктів критичної інфраструктури. Більшість ситуацій, спричинених загрозами, аваріями та катастрофами є типовими. Методики подолання їх наслідків також можуть бути універсальним і використовуватись для критичних інфраструктур різних секторів. Тому, для

цілей дослідження запропоновано використання такого підходу до класифікації об'єктів, як RFM-аналіз [71].

RFM-аналіз дозволяє розподілити кожен з сукупностей об'єктів критичної інфраструктури за характеристиками, наведеними в таблиці 1.1, відомчою підпорядкованістю та критеріями критичності (згідно чинного законодавства), тощо на декілька груп. У проведеному дослідженні, експериментальним шляхом на основі матеріалів окремих об'єктів критичної інфраструктури, визначено, що оптимальною кількістю груп є п'ять. Групи побудовані таким чином, що вони мають однакову кількість об'єктів або включають однакову кількість подій, які є заходами по забезпеченню захисту цих об'єктів або подіями, за яких відбулась кібератака. Наприклад, п'ять груп по R, де $R=5$ означає, що подія відбулась зовсім недавно, п'ять груп по F, де $F=5$ означає, що подія відбувалась часто, п'ять груп по M, де $M=5$ означає, що подія відбулась зовсім недавно. Комбінація груп дає RFM бал, який можна порівнювати між собою. Наприклад, для групи що має RFM-бал 555, можна дати наступний вербальний опис – події щодо цієї групи відбувалися часто і нещодавно, об'єкти критичної інфраструктури, що входять до цієї групи належать до I та II груп критичності, є значні ризики негативних наслідків. Використання класичного RFM-аналізу передбачає три виміри, насамперед тому, що методика походить з CRM-систем, які передбачають представлення результатів в трьохвимірному просторі. Комбінація груп дозволяє розрахувати RFM-бал наступним чином: шляхом квантильного перетворення змінні зводяться до категоріальних, потім одержані бали конкатенуються і таким чином одержується RFM-бал. Далі розраховуються перший та третій квартилі та виконується профілювання сегментів відповідно до наборів даних, що розподілялись на сегменти («високий», «середній», «низький»). Застосування даної методики забезпечить співставність показників для об'єктів, що належать до різних секторів критичної інфраструктури, дозволить визначити наявність синергетичного ефекту від несанкціонованого втручання чи настання кризової ситуації, спрогнозувати ефективність системи захисту критичної інфраструктури в цілому та її окремих об'єктів в секторі національної економіки,

регіоні, окремих містах та районах. Це створить передумови для розроблення уніфікованих рекомендацій як стосовно забезпечення захисту об'єктів критичної інфраструктури, так і для реагування на події, що відбуваються під час критичної (кризової ситуації) або атаки. Практичне застосування методики RFM-аналізу в оцінках рівня захищеності критичної інфраструктури дозволить виділити групи об'єктів одного сектору критичної інфраструктури чи адміністративно-територіальній одиниці, що вирізняються рівнем їх рівень критичності. В результаті, для кожної групи об'єктів, можна розробити стратегію, адаптовану до їх потреб та умов функціонування. Саме RFM-аналіз дозволить не тільки ранжувати окремі об'єкти за їх захищеності, а й сегментувати інформаційну базу критичної інфраструктури, при чому сегменти мають чітко виражені межі, що дозволить прогнозувати стан системи захисту критичної інфраструктури на майбутнє. Головною перевагою застосування RFM-аналізу є можливість використання багатовимірного аналізу даних, виявлення зв'язків між окремими факторами середовища, в якому функціонує критична інфраструктура. До недоліків класичного RFM-підходу можна віднести складність вербального інтерпретування, тому що при розбитті кожного виміру R, F, M, навіть на п'ять категорій, створюється 125 груп, витрати часу на вербальне пояснення кожної є досить значними.

В процесі дослідження розглянуто сукупність об'єктів критичної інфраструктури, розташованих в декількох адміністративно-територіальних одиницях регіону, згрупованих за секторальною ознакою (один із варіантів групування) та ступенем критичності (другий варіант).

З огляду на особливості критичних інфраструктур різної секторальної приналежності, запропоновано модифікувати класичну методику RFM-аналізу. Суть модифікації полягає в наступному:

а) для обчислення RFM-балу використано не три виміри, що описують подію, а довільну кількість.

б) замість трьохзначного RFM-коду використано адитивний інтегральний RFM-бал,

в) застосовано поділ сукупності на невелику кількість сегментів, наприклад трьох, замість декількох десятків та навіть сотень.

Запропонований у процесі дослідження алгоритм використання RFM-аналізу був апробований у дослідженні соціально-економічних показників, показавши прийнятні результати. В даному дослідженні він був модифікований і включає наступні кроки:

1. Виконано квантильне перетворення стосовно до кожної чисельної змінної з метою створення категоріальних змінних. Кожна змінна може приймати значення від 1 до 5.

2. Обчислено адитивний інтегральний RFM – бал за формулою (6.7):

$$RFM_value = \sum_{i=1}^n X_i_value, \quad (6.7)$$

3. Визначено перший та третій квартилі для адитивного інтегрального RFM-балу.

4. Об'єкти аналізу віднесено до відповідного кластеру за правилами: кластер “низький рівень”, якщо адитивний інтегральний RFM-бал менший за значення першого квартиля, кластер “середній рівень”, якщо адитивний інтегральний RFM-бал набуває значення між першим та третім квартилями, кластер “високий рівень”, якщо адитивний інтегральний RFM-бал більший за значення третього квартиля.

Пропонований підхід подолання невизначеностей, зокрема на етапі підготовки даних до аналізу та під час попередньої побудови моделей, передбачає застосування зокрема, статистичних та ймовірнісних підходів, моделювання досліджуваних процесів, альтернативні методи оцінювання параметрів моделей, тощо. Адже важливим є вибір оптимального інструментарію підтримки прийняття рішень, особливо, в умовах невизначеності та обмеженості в часі за різних сценаріїв розвитку подій та горизонту прогнозування, застосування уніфікованих підходів щодо розроблення рекомендацій по забезпеченню кіберзахисту, інженерного захисту, тощо, груп типових об'єктів критичної інфраструктури. Використання даного підходу

також дозволить оцінити ефективність заходів по забезпеченню захисту критичної інфраструктури. У поєднанні з методами математичного моделювання, інтелектуального аналізу даних, штучного інтелекту, запропонована методика класифікації об'єктів та подій в управлінні захистом критичної інфраструктури, дозволить оптимізувати та прискорити процес прийняття рішень в управлінні об'єктом критичної інфраструктури, зокрема у критичній (або кризовій) ситуації, варіанти рішення будуть генеровані автоматично. Крім того, поглиблене інформаційне дослідження функціонування об'єктів не лише допоможе виявити «вузькі місця» в організації захисту критичної інфраструктури, а й у випадку системної невизначеності, коли не можливо зібрати всю інформацію, необхідну для прийняття рішення, можна скористатись методикою опрацювання подібних процесів чи систем.

Перевагою використання запропонованої методики опрацювання вхідних даних є її простота та гнучкість. Пропоновані підходи можна використовувати як для аналізу неповних даних, так і для класифікації об'єктів, щоб забезпечити обґрунтованість віднесення певного об'єкта (або групи) до групи у реєстрі об'єктів критичної інфраструктури. Крім того, класифікація об'єктів на етапі попередньої обробки даних може бути реалізована й сумісно з математичними моделями та їх ансамблями, різними інструментами інтелектуального аналізу даних, тощо.

6.4 Застосування ансамблів моделей для оптимізації виробництва регіону

Задача задачу пошуку оптимальних співвідношень виробництва різних видів продукції АПК та її реалізації її на світовому та внутрішньому ринках може бути сформульована наступним чином. Альтернативами першого рівня $A = \{a_1, a_2, \dots, a_n\}$ є потенційні імпортери української аграрної продукції. Множина альтернатив першого рівня оцінюється на множині критеріїв $C = C_1UD$. Альтернативами другого рівня $B = \{b_{i1}, b_{i2}, \dots, b_{ik}\}$ є види продукції, що будуть

реалізовані, обсяги реалізації та експортери конкретних видів сільськогосподарської продукції. Підмножина критеріїв $C_1 = \{c_1, c_2, \dots, c_p\}$, використовується для оцінки альтернатив виключно першого рівня, а множина критеріїв $D = \{d_1, d_2, \dots, d_s\}$, – як для першого, так і для другого рівнів. Кожному елементу множини A («покупцю») ставиться у відповідність множина альтернатив другого рівня (обсяги реалізації кожного виду сільськогосподарської продукції за окремим напрямком). Необхідно ранжувати альтернативи множини A враховуючи багатокритеріальні оцінки альтернатив множини B . Складністю даної задачі є те, що на першому рівні для вибору альтернатив використовуються переважно якісні ознаки, а на другому – кількісні. Крім того, враховуючи те, що у процесі відбору альтернатив можуть брати участь як вітчизняні, так і закордонні експерти, їх набір, кількість та характеристики альтернатив можуть суттєво відрізнятися.

Методика використання засобів текстової аналітики, запропонована у даній роботі, призначена для відбору та аналізу альтернатив, а також визначення цільової установки дослідження передбачає послідовну реалізацію набору кроків, які охоплюють як попередню обробку текстової інформації, формування набору правил для аналізу, відбір найбільш значимих критеріїв та цілей, так і візуалізацію одержаних результатів.

Схема реалізації методика та використання засобів текстової аналітики можна представити у вигляді наступної послідовності кроків.

1. Сформувати цілі та критерії відбору альтернатив, визначивши при цьому рівні аналізу та згрупувати альтернативи за рівнями аналізу.
2. Сформувати множину альтернатив першого рівня
3. Сформувати множину альтернатив другого рівня
4. Сформулювати множину оцінок альтернатив другого рівня за найбільш значимими критеріями.
5. Обрати метод багатокритеріального аналізу. Оскільки поточна ситуація в економіці України характеризується наявністю невизначеностей різного

характеру, застосовується багатокритеріальна функція корисності альтернатив [12, с. 122] (4):

$$F(y) = \sum_{i=1}^n k_i f_i(y_i) \quad (6.8)$$

6. Побудувати однокритеріальні функції для оцінки корисності альтернатив, що відповідають критеріям для оцінки альтернатив першого та другого рівнів. При побудові функцій слід враховувати ризики для кожної оцінки корисності.

7. Виконати оцінку шкалюючих коефіцієнтів

8. Виконати оцінку альтернатив другого рівня за допомогою функції корисності.

9. Сформувати множину з найкращих альтернатив другого рівня, відповідно до множини варіантів переваг особи, що приймає рішення

10. Оцінити альтернативи першого рівня за множиною критеріїв, сформованої виключно для першого рівня. За критеріями корисності та цінності для оцінки альтернатив першого і другого рівнів, на основі множини альтернатив, сформованої на кроці 9, відібрати набори альтернатив другого рівня.

11. Виконати попарні порівняння критеріїв із множини, сформованої на кроці 10.

12. Виконати порівняння альтернатив першого рівня за критеріями, представленими у множині оцінки критеріїв для першого та другого рівнів та виконати переведення оцінок корисності в оцінки відносної важливості.

13. Виконати парні порівняння альтернатив першого рівня за критеріями множини оцінки альтернатив виключно першого рівня.

14. Виконати багатокритеріальну оцінку альтернатив першого рівня.

15. Вибрати найкращу альтернативу першого рівня на основі аналізу відносної стійкості розв'язків наборів кращих альтернатив, відібраних на кроці 10.

На етапах 1-4 запропоновано забезпечити автоматизоване опрацювання значних обсягів переважно тестової інформації за допомогою таких потужних сучасних інструментів – сімейство програмних засобів SAS Textual Analytics. Дані програмні засоби мають широкий спектр можливостей для аналізу структурованих та неструктурованих даних: це і SAS Web Crawler, SAS Search and Indexing, SAS Enterprise Content Categorization, SAS Ontology Management, SAS Text Miner та SAS Sentiment Analysis Studio.

В ході дослідження (етапи 1-4) дані для відбору альтернатив першого та другого рівнів пропонується видобувати з веб-простору, використовуючи SAS Web Crawler – програмне забезпечення, що сканує сторінки мережі Інтернет та використовує вбудовану структуру посилань, за допомогою якої веб-сторінки пов'язані одна з одною. Веб-краулеру необхідно вказати посилання на сторінку, з якої необхідно розпочати сканування. Цю веб-сторінку ще називають точкою входу. У даному дослідженні точкою входу переважно є посилання безпосередньо на розділ новин певного веб-сайту. Звичайний веб-краулер починає свою роботу з дослідження точки входу і сканує розміщені на ній посилання, котрі, у свою чергу, посилаються на інші веб-сторінки. Віднайдені таким чином сторінки також скануються на наявність посилань на інші сторінки. Глибину такого пошуку можна обмежити шляхом встановлення значень певних параметрів. У ході парсингу веб-сторінок веб-краулер вилучає з них необхідний контент і зберігає його у вигляді файлів-документів.

Для визначення емоційного забарвлення текстової інформації, запропоновано побудувати множини правил для відбору позитивних та негативних вподобань для кожної частини мови.

Для проведення дослідження, передбаченого на кроці 4 запропоновано використовувати інструменти SAS Sentiment Analysis Studio, які дозволяють аналізувати тональність тексту на основі побудованих правил, будувати моделі, в основу яких покладені правила та гібридні моделі (на основі статистичних моделей та моделей на основі правил).

Дослідження виконано на матеріалах щодо виробництва та реалізації основної експортної культури АПК України – кукурудзи. Перевірка працездатності та адекватності розробленої методики виконана в ході аналізу Інтернет-джерел текстів щодо тематики виробництва в Україні окремих сільськогосподарських культур та пріоритетних напрямків їх реалізації на внутрішньому та світовому ринках для формування альтернатив першого та другого рівня задачі дворівневого багатокритеріального аналізу альтернатив у системі підтримки прийняття рішень розвитку агропромислового виробництва України.

Для аналізу корпусу текстів було використано SAS Textual Analytics Suit, що складається з SAS Content Categorization та SAS Sentimental Analysis Studio за допомогою яких була розроблена таксономія із 776 правил. Одночасно з автоматизованим дослідженням виробництва та ринку кукурудзи, авторами виконувалось дослідження і без використання засобів текстової аналітики. Порівнюючи трудозатрати на обробку даних за двома способами, слід відзначити, що використання інструментів тестової аналітики компанії SAS значно знизило трудомісткість одержання експертного висновку: було витрачено 40 людино-годин проти 900 людино-годин без використання вказаних інструментів. Значними є обсяги опрацьованої інформації (табл. 6.8).

Таблиця 6.8

Перелік Інтернет-джерел, що опрацьовані в ході дослідження

Назва Інтернет-ресурсу	Кількість завантажених статей
http://latifundist.com/blog	5649
http://agroua.net/news/	1324
http://obozrevatel.com/blogs	2741
http://blog.liga.net	11008
http://blogs.lb.ua	3538
http://agroportal.ua/views/blogs	1038
http://www.agro24.com.ua/blog	116
http://blogs.pravda.com.ua/	23180

На рис. 6.16 представлений приклад побудови таксономії в SAS Content Categorization Studio.

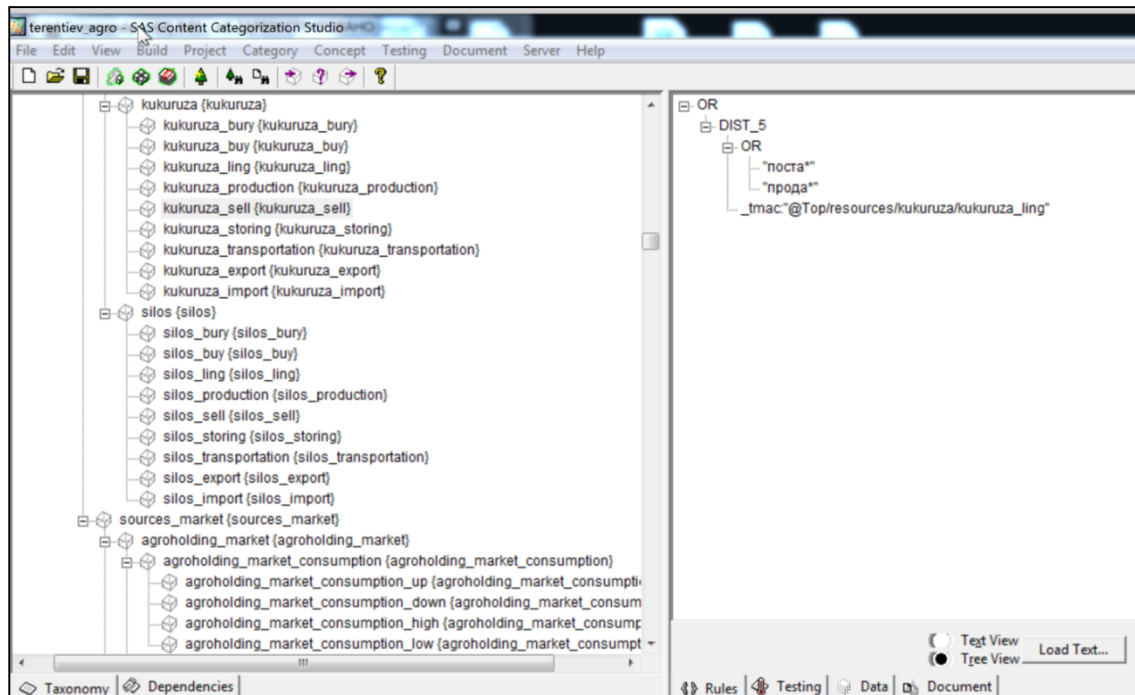


Рис. 6.16. Результати побудови правил у програмі SAS Content Categorization Studio

На рис. 6.16 відображено, як правило (на правій панелі) дозволяє на дистанції у 5 слів виконувати пошук словосполучень типу «постачання кукурудзи», «продаж кукурудзи».

Для зручності побудови звітів отриманий набір даних, який по суті являє собою багатомірну OLAP-базу даних, було розкладено за чотирма рівнями.

Синтез правил щодо аналізу досліджуваної системи на прикладі домена «сільське господарство» через виявлення станів «високий, низький, зростаючий або спадаючий рівень потенційно негативного або позитивного показника» із додатковим вилученням фактуальної словникової інформації (словник-категоризатор країн, ієрархія понять з галузі сільського господарства з класифікатору КВЕД – табл. 6.9).

Таблиця 6.9

Назви груп КВЕД для генерації правил категоризатору

Тип	H1	H2	H3	H4	Текст	Ключові слова
LI	А. Сільське господарство, лісове господарство та рибне господарство	01. Сільське господарство, мисливство та надання пов'язаних із ними послуг	01.1. Вирощування однорічних і дворічних культур	01.11. Вирощування зернових культур (крім рису), бобових культур і насіння олійних культур	Вирощування зернових культур, таких як:	Вирощування, вирощування зернових культур, зернова культура, культура
LI	А. Сільське господарство, лісове господарство та рибне господарство	01. Сільське господарство, мисливство та надання пов'язаних із ними послуг	01.1. Вирощування однорічних і дворічних культур	01.11. Вирощування зернових культур (крім рису), бобових культур і насіння олійних культур	пшениця	пшениця
LI	А. Сільське господарство, лісове господарство та рибне господарство	01. Сільське господарство, мисливство та надання пов'язаних із ними послуг	01.1. Вирощування однорічних і дворічних культур	01.11. Вирощування зернових культур (крім рису), бобових культур і насіння олійних культур	кукурудза (на зерно)	зерно кукурудза
LI	А. Сільське господарство, лісове господарство та рибне господарство	01. Сільське господарство, мисливство та надання пов'язаних із ними послуг	01.1. Вирощування однорічних і дворічних культур	01.11. Вирощування зернових культур (крім рису), бобових культур і насіння олійних культур	сорго	сорго
LI	А. Сільське господарство, лісове господарство та рибне господарство	01. Сільське господарство, мисливство та надання пов'язаних із ними послуг	01.1. Вирощування однорічних і дворічних культур	01.11. Вирощування зернових культур (крім рису), бобових культур і насіння олійних культур	ячмінь	ячмінь

Предметний домен S_0 «сільське господарство» має наступні складові S_{11} та S_{12} : <‘культури’, ‘ринки’> згідно моделі опису складної ієрархічної системи.

Рівні S_{1i} складаються у свою чергу з інших рівнів S_{2j} , наприклад: $S_{11} = \langle \text{‘кукурудза’, ‘пшениця’, ‘ячмінь’, ‘зерно’, ‘гречка’, ‘соняшник’, ‘ріпак’} \rangle$.

Кожне джерело рівня S_{11} має над собою функціональні залежності, що перетворюють його кількісно чи якісно: $\phi_k(S_{1j}) \in \langle \text{‘покупка’, ‘продаж’, ‘виращування’, ‘селекція’, ‘транспортування’, ‘зберігання’, ‘переробка’, ‘використання’} \rangle$ та впливають на зовнішні параметри системи S_0 .

Наступним кроком є синтез схеми, у якій за відповідними шаблонами зв'язуються показники з економічного словника (kpi) та рівні ієрархії предметного домену. Для визначення трендів до застосованих вище показників додані типові показники з економічного тезаурусу:

$$kpi_trends = kpi \times lvl \cap kpi \times dir \quad (6.9)$$

де $kpi = \langle arg_j \rangle$, $arg_j \in \langle \text{‘споживання’, ‘вартість’, ‘дефіцит’, ‘попит’, ‘знижка’, ‘надлишок’, ‘інвестиції’, ‘вихід’, ‘ціна’, ‘квота’, ‘ризик’, ‘частка’, ‘ринкова вартість’, ‘субсидія’, ‘поставка’, ‘тариф’, ‘ставка податку’, ‘обсяг’} \rangle$ та інші, $lvl = \langle arg_k \rangle$, $arg_k \in \langle \text{‘великий’, ‘низький’} \rangle$, $dir = \langle arg_l \rangle$, $arg_l \in \langle \text{‘зріст’, ‘спад’} \rangle$.

В результаті проведеного дослідження одержано наступні результати. Виявлено, що найбільш обговорюваними темами щодо тенденцій розвитку світового ринку аграрної продукції є наявність генно-модифікованих організмів (7,42%), виробництво біодизеля (34,68%), проблеми вирощування кукурудзи (55,87%).

Щодо відбору альтернатив другого рівня, пов'язаних із учасниками ринку кукурудзи, то загалом на блогах, форумах та сайтах новин, присвячених сільському господарству, найчастіше згадують про Україну – 21%, Росію – 17%, Європейський Союз – 17%, США – 10%, Польщу – 8,5% та інші країни – 26,5%. Найбільш важливими є альтернативи нарощування обсягів виробництва кукурудзи – 56%, виробництво біопалива – 34%, генетично модифіковані організми – 7% та інші – 4%.

В ході аналізу альтернатив значення цін та розвитку ринку сільськогосподарських культур, було виявлено, що розподіл пріоритетів змінювався протягом 2014-2016 маркетингових років. Якщо у 2015-му році

структура альтернатив була наступною: ринкові розрахунки - 30%, інвестиційна діяльність – 15%, ринковий попит – 10%, ринкова пропозиція – 10%, розподіл ринку – 10%, ефективність вирощування культур – 5%. У 2016 році структура дещо відрізняється експертів більше цікавили питання про ринкові розрахунки (25%) та інвестиційну діяльність (20%).

Щодо емоційного забарвлення досліджуваної інформації, то слід відзначити зростання негативної тенденції щодо «квоти на продаж продукції» та «генно-модифіковані організми», позитивної – «кукурудза» та «Китай».

При порівнянні альтернатив, враховуючи наявність ризиків, використано метод аналізу ієрархій. Для переведення оцінок корисності у оцінки порівнюваної важливості використовувалися коефіцієнт переведення [12] (5):

$$K = \frac{c_{\max} - 1}{u_{\max} - u_{\min}}, \quad (6.10)$$

де $u_{\max}$ та $u_{\min}$ відповідно, максимальна та мінімальна корисність альтернатив за певним критерієм,

$c_{\max}$ – максимальна оцінка відносної важливості критеріїв.

Максимальна оцінка важливості альтернатив, визначається відповідно до шкали оцінок особою, що приймає рішення [12, с. 124].

Основними перевагами застосування текстової аналітики та дослідження емоційного забарвлення у процесі відбору альтернатив є можливість опрацювання значних обсягів неструктурованої інформації, що накопичена в різних джерелах, виявивши при цьому якнайбільшу кількість альтернатив за мінімальних трудовитрат особи, що приймає рішення значно знизивши при цьому суб'єктивізм експертних оцінок.

При прийнятті рішення перевага віддається альтернативі з максимальним значенням математичного очікування та мінімальним стандартним відхиленням ваг. Ці значення розраховуються для кожного сценарію окремо.

Маючи кількісні характеристики альтернатив сценаріїв, на наступному рівні аналізу, особа що приймає рішення має змогу візуалізувати профіль альтернативи сценарію у певний момент часу часового проміжку, порівняти між

собою часові перерізи альтернатив одного або декількох сценаріїв у визначений момент часу. Кожний фактор, опрацьований за запропонованою методикою, описується певним інтегральним показником, до складу якого входять згруповані та впорядковані за важливістю для даного фактору лінгвістичні змінні у вигляді вербальних описів можливих варіантів розвитку подій. Це дає змогу формалізувати нечіткі уявлення щодо розвитку досліджуваного соціально-економічного процесу в рамках певного сценарію, побудованого на основі тенденцій, виявлених експертами.

Отже, використання SAS Textual Analytics дозволило знизити вплив людського фактору на вибір альтернатив та значно пришвидшити опрацювання значних обсягів неструктурованої інформації з Інтернет-джерел та сформувані найбільш повне уявлення щодо розвитку ситуації, наявності ризиків та ознак кризових явищ, створити передумови для дослідження багатofакторних причинно-наслідкових залежностей, що можуть виникати серед різномірної та неструктурованої інформації, джерелом якої є відомості, одержувані з офіційних та неофіційних джерел.

6.5 Моделювання еколого-економічних показників

Враховуючи існування тісного зв'язку результативності сільського господарства із станом та динамікою біоресурсів і природокористування, потребує подальшого удосконалення існуюча система його інформаційно-аналітичного забезпечення, орієнтована на опрацювання переважно ретроспективних статистичних даних щодо обсягів виробництва продукції та основних економічних процесів. Фактично при цьому не використовуються геоінформаційні технології, інформаційно-аналітичні системи, призначені для прогнозування метеорологічних умов, продуктивності тварин, формування оптимального раціону відгодівлі худоби, формування технологічних карт вирощування сільськогосподарських культур на основі технологій точного землеробства та раціонального використання природних ресурсів.

Сьогодні для оцінювання урожайності сільськогосподарських культур використовуються різні підходи, серед яких переважають статистичні методи, у

яких використовується припущення щодо існування залежності між характеристиками навколишнього середовища, внесенням добрив та урожайністю.

Біофізичні моделі, у яких в якості вхідних змінних використовуються дані дистанційного зондування з метою калібрування моделі та її налаштування, застосовуються набагато рідше. Основним недоліком таких моделей є необхідність задавання численних специфічних параметрів для кожної культури: характеристик ґрунту, методів обробки культур, агрометеорологічних даних, строків посіву. Все частіше для побудови моделей, що описують залежність урожайності сільськогосподарських культур від обсягів внесених поживних речовин, стану ґрунтів, впливу факторів навколишнього середовища, застосовують дані, одержані в результаті дистанційного зондування [10]. Ці методи, як правило, прості, а їх реалізація не вимагає задавання значної кількості параметрів. Однак, необхідно враховувати, що процеси, які описують урожайність основних сільськогосподарських культур, як правило, є нестационарними, а показники урожайності необхідно опрацьовувати для значної кількості господарств певної ґрунтово-кліматичної зони.

Статистичні методи прогнозування врожайності на основі оцінок фізичних факторів середовища ґрунтуються на припущенні про існування простої залежності між характеристиками навколишнього середовища та врожайністю. Крім того, враховуючи визначальну роль стану формування біомаси рослин у певні фази їх розвитку, для моделювання урожайності сільськогосподарських культур застосовують кількісні показники фотосинтетичної активності біомаси – вегетаційні індекси. Один із них – NDVI (Normalized Difference Vegetation Index) – нормалізований відносний індекс вегетації – простий кількісний показник кількості фотосинтетичної активної біомаси (так званий вегетаційний індекс). В цілому, головною перевагою NDVI є простота його отримання: для обчислення індексу не потрібно ніяких додаткових даних і методик, окрім безпосередньо самої космічної зйомки і знання її параметрів. Динаміка сезонної зміни NDVI дає можливість оцінювати стан рослин у різних фазах розвитку.

Методи, запропоновані для розв'язування задач даного класу рідко застосовуються на практиці внаслідок значної трудомісткості необхідних наземних вимірювань і складності досліджуваних залежностей [6].

У виконаному дослідженні інформаційною базою є статистичні дані про урожайність основних сільськогосподарських культур за 2007-2013 рр., вирощуваних у зоні Лісостепу на полях підприємств, що входять до складу одного з українських холдингів, а також дані щодо впливу фізичних факторів середовища та вегетаційні NDVI індекси основних сільськогосподарських культур.

Побудова моделей за експериментальними даними має істотну теоретичну та практичну основу [1, 6, 7, 10, 11], зокрема, широкий арсенал методів і засобів структурної та параметричної ідентифікації. Крім того, ці засоби та методи можуть мати як універсальний характер, так і бути певною мірою орієнтованими на конкретні предметні галузі, різний рівень формалізації задачі. Одними з найбільш широко застосовуваних для аналізу експериментальних даних в економіці, фінансах, соціології, біології, психології, медицині, техніці тощо є методи регресійного аналізу [6, **Ошибка! Источник ссылки не найден.**,12].

Парні (прості) лінійні регресійні моделі встановлюють лінійну залежність між двома змінними [12], наприклад залежність урожайності зернових від кількості внесених добрив, витрат на обробіток ґрунту на полях сівозміни та обсягів реалізованої продукції.

У загальному випадку парна вибіркова регресійна модель [13, с. 244] має вигляд:

$$Z = a_0 + a_1X + E, \quad (6.11)$$

де Z – вектор значень залежної змінної, $Z = [z_1, z_2, \dots, z_n]$; X – вектор значень незалежної змінної, $X = [x_1, x_2, \dots, x_n]$; a_0, a_1 – невідомі параметри регресійної моделі; E – вектор значень випадкових величин $E = [e_1, e_2, \dots, e_n]$, поява яких у моделі зумовлена наявністю випадкових збурень, недосконалою структурою моделі, похибками вимірів та обчислень оцінок параметрів.

При великій кількості вільних (екзогенних) змінних виникає проблема нестійкого оцінювання параметрів моделі, а тому застосовують метод найменших кутів [14] (LARS, least-angle regression) – алгоритм відбору ознак в задачах лінійної регресії та метод часткових найменших квадратів (PLS, partial least squares) [15]. Алгоритм LARS схожий на алгоритм покрокової регресії. Відмінність полягає у тому, що алгоритм LARS, замість послідовного додавання вільних змінних, на кожному кроці змінює їх вагові коефіцієнти.

Метод PLS [15] – статистичний метод, який має деяке відношення до регресії за методом головних компонент, замість знаходження гіперплощин з мінімальною дисперсією між вихідною і незалежними змінними, він знаходить модель лінійної регресії шляхом проектування прогнозованих і спостережуваних змінних у новому просторі. Оскільки обидва набори даних X і Y переноситься у новий простір, то сімейство PLS-методів відоме як моделі білінійних факторів.

Модель регресії, побудована за методом PLS, особливо підходить у випадку, коли матриця предикторів має більше змінних, ніж спостережень, і коли у матриці вимірів X є мультиколінеарність.

Серед сучасних методів розв’язання задач у багатовимірному просторі також необхідно виділити методи градієнтного бустингу [16]. Задача бустингу у загальному випадку формулюється таким чином [16]. Нехай, заданий багатовимірний простір X з простором міток Y , є навчальною вибіркою $\{x_i\}_{i=1}^N$, де $x \in X$. Задача зводиться до пошуку пари оптимальних параметрів $\{a_m, b_m\}$ для класифікатора F_m довжини m :

$$F_m(x) = F_{m-1}(x) + b_m h(x; a_m), b_m \in R, a_m \in A. \quad (6.12)$$

Оптимальність розв’язку визначається відповідно до принципу явної максимізації відступів, що передбачає введення деякої функції втрат $L(y_i, F_m(x_i))$, $i = \overline{1, N}$, яка показує наскільки змодельований результат відрізняється від правильної відповіді y_i . Потім мінімізується функціонал

похибки Q (3) [16]:

$$Q_i = \sum_{i=1}^N L(y_i, F_m(x_i)) \rightarrow \min \quad (6.13)$$

Для розв'язання задачі мінімізації функціоналу виконують один крок алгоритму градієнтного спуску. У точці, для якої шукають оптимальний приріст, розраховують градієнт функціоналу похибки [16]:

$$\nabla Q_i = \left[\frac{\partial Q}{\partial F_{m-1}}(x_i) \right]_{i=1}^N = \left[\frac{\partial \left(\sum_{i=1}^N L(y_i, F_{m-1}) \right)}{\partial F_{m-1}}(x_i) \right]_{i=1}^N = \left[\frac{\partial L(y_i, F_{m-1})}{\partial F_{m-1}}(x_i) \right]_{i=1}^N \quad (6.14)$$

Далі, використовуючи відповідні алгоритми навчання, знаходять пари оптимальних параметрів $\{a_m, b_m\}$: для визначення a_m використовується базовий алгоритм навчання, а для b_m – лінійний пошук.

Алгоритм градієнтного бустингу застосовується у задачах регресійного моделювання. Це, наприклад, метод LS-boosting [16], коли функція втрат має вигляд:

$$L(y, F) = \frac{(y - F)^2}{2}, \quad (6.15)$$

а градієнт функціоналу помилки (∇Q) в алгоритмі градієнтного бустингу набуває такого вигляду:

$$\nabla Q_i = y_i - F_{m-1}(x_i), \quad (6.16)$$

де y_i – значення міток для кожного об'єкта.

Іншим варіантом є алгоритм мінімізації середнього модуля відхилення (LAD-bosting) [16]. Функція втрат має вигляд:

$$L(y, F) = |y - F|. \quad (6.17)$$

Тоді ∇Q набуде вигляду:

$$\nabla Q_i = \text{sign}(y_i - F_{m-1}(x_i)). \quad (6.18)$$

В цьому випадку b_m розраховується через зважене медіанне значення вибірки, тобто виконується задача пошуку порядкової статистики в масиві.

У випадках наявності шуму вимірів та викидів в даних кращі результати можна одержати за допомогою процедури Tree-boosting, оскільки в даному випадку ∇Q_i набуває значень $\{+1, -1\}$, а оновлення параметрів відбувається за рахунок медіанних значень. Ще одним методом, що добре зарекомендував себе у задачах з сильно зашумленими даними, є метод М-бустингу [16]. Цей метод використовує не квадратичну функцію втрат, а функцію втрат Хубера, яка має вигляд:

$$L(y, F) = \begin{cases} \frac{1}{2}(y - F)^2, & |y - F| \leq \sigma \\ \sigma(|y - F| - \frac{\sigma}{2}), & |y - F| > \sigma \end{cases} \quad (6.19)$$

Це дозволяє використовувати метод найменших квадратів для незначних похибок та функцію втрат – для великих. Градієнт у цьому випадку має вигляд (10):

$$\nabla Q_i = \begin{cases} y_i - F_{m-1}(x_i), & |y_i - F_{m-1}(x_i)| \leq \sigma \\ \sigma - \text{sign}(y_i - F_{m-1}(x_i)), & |y_i - F_{m-1}(x_i)| > \sigma \end{cases} \quad (6.20)$$

Оскільки розподіл похибок на кожній ітерації бустингу різний, то на кожній ітерації підбирається нове значення σ_m (?), виходячи з різних міркувань, за формулою (6.21):

$$\sigma_m = \text{quantile}_\alpha \left\{ |y_i - F_{m-1}(x_i)| \right\}_{i=1}^N \quad (6.21)$$

Оскільки процеси, що мають місце у сільськогосподарському виробництві, характеризуються множиною сильно взаємодіючих змінних, а крім того, є необхідність формувати структури даних, які можна використати для

розв'язання інших задач обробки сільськогосподарських даних, то виникає необхідність використання композиція різних методів.

Враховуючи особливості сільськогосподарського виробництва, зокрема вплив попередника у сівозміні, неоднорідність ґрунтів, особливості технологій вирощування тощо, при оцінюванні впливу різних факторів на зміну урожайності різних сільськогосподарських культур, виникає проблема виявлення прихованих закономірностей в оброблюваних даних. Для вирішення даної задачі якісні прогнози можна одержати застосовуючи такий різновид дерев рішень, як дерева регресії. Побудова регресійних моделей із застосуванням дерев рішень відбувається у два етапи: спочатку, для відбору найбільш значущих змінних будується дерево рішень, потім – модель лінійної регресії, у якій використано одержані на попередньому етапі набори найбільш значимих змінних. Сьогодні розроблено та реалізовано програмне забезпечення для побудови дерев рішень за алгоритмами C4.5, C5.0, CART, CHAID, CN2, SPRINT, ID3 та ін. Перевагами дерев рішень є наочність, швидке навчання, висока точність прогнозу, можливість розробки правил у випадках, коли знання складно формалізувати.

Формалізм побудови узагальнених графічних моделей поєднує у собі багато методів статистичного моделювання, таких як факторний аналіз, аналіз розподілів, моделі сумішей розподілів, приховані марковські моделі, фільтри Калмана, моделі Айзинга та деякі інших. Перевагою такого підходу є те, що методи дослідження процесів та обробки даних, розроблені в одній області, можуть бути успішно перенесені в інші **[Ошибка! Источник ссылки не найден.]**.

Мережа Байєса представляє собою пару $\langle G, B \rangle$ [17], у якій перша компонента G – це спрямований нециклічний граф, що відповідає випадковим змінним і записується як множина умов незалежності: кожна змінна незалежна від її батьків в G . Друга компонента пари B – це множина параметрів, що визначають мережу.

Компонента B містить параметри $\Theta_{X^{(i)}|\rho\alpha(X^{(i)})} = P(X^{(i)}|\rho\alpha(X^{(i)}))$ для кожного можливого значення $x^{(i)} \in X^{(i)}$ та $\rho\alpha(X^{(i)}) \in P\alpha(X^{(i)})$, де $P\alpha(X^{(i)})$ позначає множину батьків змінної $X^{(i)} \in G$. Кожна змінна $X^{(i)} \in G$ представляється у вигляді вершини. Якщо розглядають більше одного графа, то для визначення батьків змінної $X^{(i)}$ в графі G використовують позначення $P\alpha^G(X^{(i)})$. Повна спільна ймовірність мережі Байєса обчислюється за формулою [17]:

$$P_B(X^{(1)}, \dots, X^{(N)}) = \prod_{i=1}^n P_B(X^{(i)} | P\alpha(X^{(i)})) \quad (6.22)$$

З математичної точки зору мережа Байєса – це модель подання існуючих і відсутніх імовірнісних залежностей. При цьому зв'язок $A \rightarrow B$ являється причинним, коли подія A – причина виникнення B , тобто існує механізм, відповідно до якого значення, прийняте A , впливає на значення, прийняте B .

Теорія побудови мереж Байєса ґрунтується на припущенні, що події є вичерпними і не перетинаються. Якщо ця умова не виконується, то результати застосування мережі будуть неконсистентними (тобто неточними). У випадку коли події вичерпні і не перетинаються, то ймовірність події E можна обчислити за допомогою умовних ймовірностей [17]:

$$P(E) = \sum_{i=1}^n P(E \cap H_i) = \sum_{i=1}^n P(E|H_i) \cdot P(H_i) \quad (6.23)$$

Використовуючи рівняння (6.23), суму перетинів події E з H можна виразити:

$$P(E \cap H_k) = P(E|H_k) \cdot P(H_k) = P(H_k|E) \cdot P(E)$$

З рівності цієї рівності знайдемо, що

$$P(H_k|E) = \frac{P(E|H_k) \cdot P(H_k)}{P(E)}$$

а із врахуванням отримаємо вираз [17]:

$$P(H_k|E) = \frac{P(E|H_k) \cdot P(H_k)}{\sum_{i=1}^n P(E|H_i) \cdot P(H_i)} \quad (6.24)$$

який представляє собою формулу Байєса. На основі цієї формули будуються такі мережі.

Тут H_k означає будь-яку гіпотезу з n можливих. Ймовірності $P(E|H_k)$ задаються експертами апіорно або розраховують за навчальними даними. Ймовірності $P(E|H_k)$ є дуже корисними, тому що, як правило, легше знайти ймовірність послідовності подій типу причина-наслідок, ніж навпаки. Значення $P(H_k)$ називають апіорними ймовірностями, вони визначають початкові ймовірності для всіх гіпотез. Переваги байєсівського методу полягають у тому, що апіорні ймовірності можна уточнювати (оновлювати) відповідно до реалій протікання процесу, що досліджується. Це дозволяє уточнювати ймовірності подій при надходженні додаткової інформації.

Враховуючи необхідність опрацювання значних обсягів статистичних даних про урожайність сільськогосподарських культур, показників NDVI та погодних факторів, для моделювання та прогнозування використано програмне забезпечення SAS Enterprise Miner [8].

Для виконання дослідження запропоновано методику, яку можна представити у вигляді такої послідовності кроків:

1. Завантаження даних.
2. Розділення даних випадковим чином на 2 підвибірки: навчальну (70%) та перевірочну (30%).
3. Побудова моделей.
4. Порівняння моделей за статистичними критеріями якості.
5. Побудова загального звіту за виконаним обчислювальним експериментом.

Виходячи з необхідності опрацювання значних обсягів даних, моделювання виконувалося за допомогою SAS Enterprise Guide. Типовим прикладом є дослідження урожайності озимої пшениці. Як видно з графіку на рис. 1, значення математичного сподівання показника урожайності врожайності відрізняється у ці роки у 1,62 рази, а дисперсія – у 6,13 рази, що підтверджує нестационарність цього процесу.

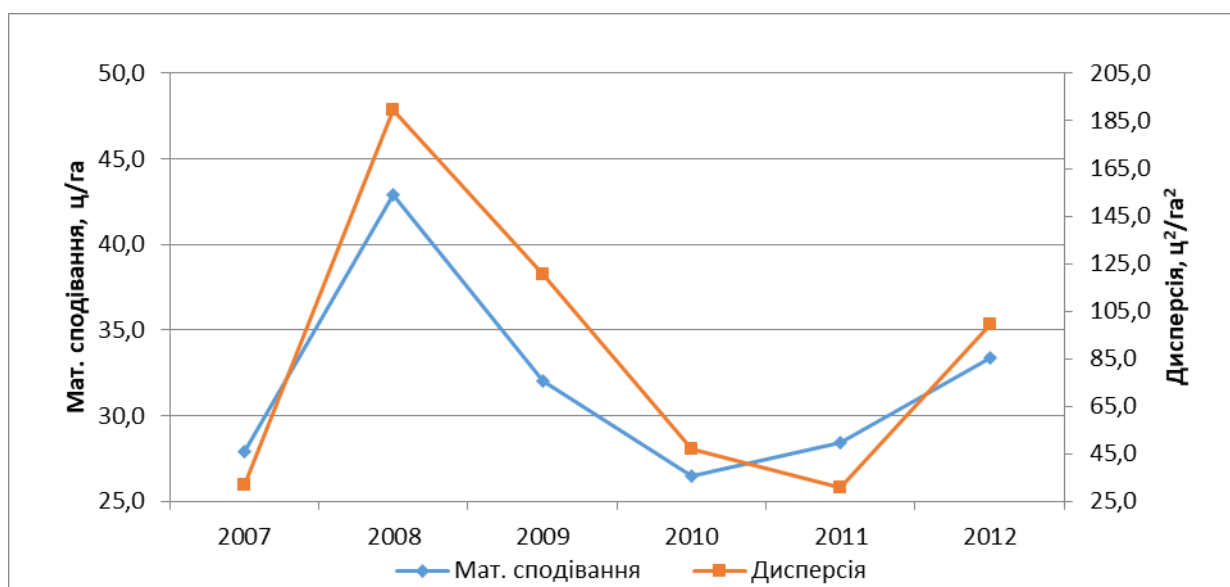


Рис.6.17. Графік умовного математичного сподівання та дисперсії урожайності озимої пшениці за 2007-2012 рр.

На рис. 6.18 зображені графіки автокореляційної та часткової автокореляційної функцій часового ряду урожайності кукурудзи (з 1991 по 2013 рр.). Як видно з рис. 6.18, характер АКФ і ЧАКФ не є постійно спадаючими. Значення АКФ і ЧАКФ є значними навіть для других різниць для першого, другого, третього, тринадцятого і вісімнадцятого лагу (а також, для набору лагів для вихідного ряду та перших різниць), що свідчить про наявність тренду другого порядку у значеннях ряду.

Вихідний ряд					Перші різниці					Другі різниці				
Autocorrelation	Partial Correlation	AC	PAC		Autocorrelation	Partial Correlation	AC	PAC		Autocorrelation	Partial Correlation	AC	PAC	
		1	0.646	0.646			1	-0.665	-0.665			1	-0.718	-0.718
		2	0.712	0.506			2	0.330	-0.203			2	0.317	-0.409
		3	0.470	-0.193			3	-0.116	0.023			3	-0.121	-0.232
		4	0.427	-0.113			4	0.144	0.239			4	0.147	0.133
		5	0.273	-0.037			5	-0.195	-0.015			5	-0.247	-0.090
		6	0.221	-0.004			6	0.263	0.133			6	0.256	-0.044
		7	0.086	-0.112			7	-0.181	0.114			7	-0.158	0.026
		8	0.058	-0.005			8	0.078	-0.018			8	0.049	-0.012
		9	-0.109	-0.192			9	0.003	0.006			9	-0.005	0.005
		10	-0.200	-0.253			10	-0.026	-0.063			10	-0.000	-0.066
		11	-0.251	0.093			11	-0.011	-0.065			11	-0.035	-0.090
		12	-0.306	0.052			12	0.042	-0.032			12	0.086	0.045
		13	-0.363	-0.191			13	-0.149	-0.233			13	-0.185	-0.231
		14	-0.332	0.054			14	0.232	0.101			14	0.245	-0.022
		15	-0.347	0.084			15	-0.234	0.011			15	-0.219	-0.041
		16	-0.277	0.003			16	0.149	0.010			16	0.182	0.097
		17	-0.335	-0.196			17	-0.189	-0.205			17	-0.213	-0.129
		18	-0.257	0.045			18	0.190	-0.073			18	0.199	-0.220
		19	-0.278	-0.030			19	-0.205	0.006			19	-0.151	-0.071
		20	-0.168	0.026			20	0.135	-0.089			20	0.071	-0.141
		21	-0.126	0.189			21	-0.095	-0.013					
		22	-0.044	-0.072										

Рис. 6.18. Показники АКФ та ЧАКФ врожайності кукурудзи в Україні з 1991 по 2013 рр.

На рис. 3 подано графік динаміки урожайності кукурудзи в Україні з 1991 по 2013 роки з накладеним трендом другого порядку. Критерії якості побудованого тренду: $R^2 = 0,86$, $R^2_{adj} = 0,84$.

У результаті дослідження отримано моделі лінійної регресії, LARS, PLS, градієнтного бустингу для прогнозування урожайності таких сільськогосподарських культур, як озима пшениця та кукурудза.

Моделювання залежності врожайності від внесених матеріалів виконано на прикладі даних про урожайність кукурудзи для двох груп полів. Побудована модель лінійної регресії за критерієм максимізації R^2 .



Рис. 6.19. Урожайність кукурудзи в Україні з 1991 по 2013 роки і накладений тренд другого порядку

Отримані результати для кукурудзи, перша група полів (6.25):

$$Productivity = N \cdot 0,01534 + P \cdot 0,06844 + 1,22805 \quad (6.25)$$

для другої групи полів (16):

$$Productivity = N \cdot 0,00807 + P \cdot 0,0834 + 4,55825 \quad (6.26)$$

Статистичні характеристики моделювання залежності урожайності кукурудзи від обсягів внесених добрив за методом лінійної регресії (друга група полів)

Критерій	RASE	RASE, %	R ²	R ² _{adj}
Значення для 1-й групи полів	0,22	2,87	0,89	0,87
Значення для 2-ї групи полів	0,05	0,84	0,88	0,86

Для оцінювання статистичних характеристик моделей у дослідженні використані показники [12, с. 201]:

SSE – сума квадратів похибок моделі:

$$SSE = \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (6.27)$$

Коефіцієнт детермінації:

$$R^2 = \frac{\text{Var}(\hat{y})}{\text{Var}(y)}, \quad (6.28)$$

Скоригований коефіцієнт детермінації R²:

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n-1}{n-k} \quad (6.29)$$

RASE (Root Average Square Error) – корінь квадратний з середньої квадратичної похибки:

$$RASE = \sqrt{\frac{1}{N-k} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (6.30)$$

MAX E (Maximum Error) – абсолютна величина максимальної похибки:

$$MAXE = \max_{i=1, \dots, N} |y_i - \hat{y}_i|, \quad \text{кількість параметрів моделі.}$$

Моделювання залежності урожайності NDVI, погодних факторів виконано на матеріалах урожайності озимої пшениці. Побудована модель лінійної регресії має вигляд:

$$\begin{aligned} \text{Productivity} = & 25,352 \cdot \text{Intercept} + 89,646 \cdot \text{ndvi_17_18} - 29,044 \cdot \text{ndvi_19_20} \\ & + 74,230 \cdot \text{ndvi_23_24} - 89,578 \cdot \text{ndvi_25_26} + 52,028 \cdot \text{ndvi_27_28} + 0,330 \cdot \text{temp_17_18} \\ & + 0,711 \cdot \text{temp_19_20} - 4,118 \cdot \text{temp_21_22} \end{aligned} \quad (6.31)$$

Регресори виду ndvi_17_18 позначають різницю між значенням нормалізованого відносного індексу вегетації (NDVI) на 18 та 17 тижнях, відповідно, temp_17_18 позначає різницю між температурою навколишнього середовища на 18 та 17 тижнях вегетації.

У таблиці 2 наведені статистичні характеристики результатів.

Таблиця 6.11

Статистичні характеристики моделювання урожайності пшениці озимої залежно від значень нормалізованого відносного індексу вегетації за методом лінійної регресії

Вибірка	Навчальна			Перевірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	5,21	977,74	13,64	5,54	459,81	11,73

Модель, побудована за методом LARS з вибором моделі за значенням ASE на перевіірочній вибірці має вигляд (6.32):

$$\begin{aligned} \text{Productivity} = & -32,616 \cdot \text{Intercept} + 89,325 \cdot \text{ndvi_17_18} - 28,519 \cdot \text{ndvi_19_20} \\ & + 3,361 \cdot \text{ndvi_21_22} + 70,514 \cdot \text{ndvi_23_24} - 89,180 \cdot \text{ndvi_25_26} \\ & + 51,874 \cdot \text{ndvi_27_28} - 11,133 \cdot \text{precip_17_18} - 1,462 \cdot \text{precip_19_20} \\ & + 0,572 \cdot \text{temp_21_22} \end{aligned} \quad (6.32)$$

Регресор precip_{17_18} – це різниця між кількістю опадів на 18 та 17 тижнях вегетації. У табл. 3 наведені статистичні характеристики результатів.

Таблиця 6.12

Статистичні характеристики моделювання урожайності пшениці озимої в залежності від впливу факторів зовнішнього середовища за методом LARS з вибором моделі по значенню ASE на перевіірочній вибірці

Вибірка	Навчальна			Перевіірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	5,21	977,73	13,64	5,57	465,68	11,74

Модель за методом LARS з вибором моделі по значенню SBC (6.33):

$$\begin{aligned} \text{Productivity} = & -2,641 \cdot \text{Intercept} + 74,067 \cdot \text{ndvi}_{17_18} + 10,496 \cdot \text{ndvi}_{27_28} \\ & - 7,323 \cdot \text{precip}_{17_18} - 0,286 \cdot \text{precip}_{19_20} - 0,932 \cdot \text{temp}_{21_22} \end{aligned} \quad (6.33)$$

Статистичні характеристики результатів наведені у таблиці 6.13.

Таблиця 6.13

Статистичні характеристики моделювання урожайності пшениці озимої під впливом факторів зовнішнього середовища за методом LARS з вибором моделі по значенню SBC

Вибірка	Навчальна			Перевіірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	5,54	1104,23	15,74	6,30	595,91	12,65

Результат за моделювання урожайності озимої пшениці за методом PLS (6.34):

$$\begin{aligned} \text{Productivity} = & -22,657 \cdot \text{Intercept} + 89,325 \cdot \text{ndvi}_{17_18} - 28,519 \cdot \\ & \text{ndvi}_{19_20} + 3,361 \cdot \text{ndvi}_{21_22} + 70,514 \cdot \text{ndvi}_{23_24} - 89,180 \cdot \text{ndvi}_{25_26} \\ & + 51,874 \cdot \text{ndvi}_{27_28} - 2,887 \cdot \text{precip}_{17_18} - 0,568 \cdot \text{precip}_{19_20} + 1,240 \cdot \\ & \text{precip}_{21_22} + 1,340 \cdot \text{temp}_{23_24} - 0,702 \cdot \text{temp}_{25_26} \end{aligned} \quad (6.34)$$

У таблиці 6.14 наведені статистичні характеристики результатів.

Таблиця 6.14

Статистичні характеристики моделювання залежності урожайності пшениці озимої від впливу факторів навколишнього середовища за методом PLS

Вибірка	Навчальна			Перевірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	5,21	977,73	13,64	5,57	465,68	11,74

Метод лінійної регресії з відбором значущих змінних за допомогою дерева рішень:

$$Productivity = 41,648 \cdot Intercept + 94,555 \cdot ndvi_{17_18} - 3,955 \cdot temp_{21_22} \quad (6.35)$$

У таблиці 6.15 наведені статистичні характеристики результатів.

Таблиця 6.15

Статистичні характеристики моделювання залежності урожайності пшениці озимої від впливу факторів навколишнього середовища за методом лінійної регресії з відбором значущих змінних за допомогою дерева рішень

Вибірка	Навчальна			Перевірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	5,54	1105,32	14,27	6,42	618,85	13,16

Для методу градієнтного бустингу в якості значимих змінних біли обрані: avg.air.temp_17_18, avg.air.temp_21_22, ndvi_17_18, ndvi_19_20, ndvi_21_22, ndvi_23_24, ndvi_25_26, ndvi_27_28.

Регресор виду avg.air.temp_17_18 – це різниця між середньою недільною температурою повітря на 18-у та 17-у тижнях вегетації.

У таблиці 7 наведені статистичні характеристики результатів.

Таблиця 6.16

Статистичні характеристики моделювання залежності урожайності
пшениці озимої від впливу факторів навколишнього середовища за
методом градієнтного бустингу

Вибірка	Навчальна			Перевірочна		
Показник	RASE	SSE	MAX E	RASE	SSE	MAX E
Значення	3,63	475,25	10,58	4,94	366,78	13,04

Regression		PLS		LARS Validation		LARS SBC		Gradient Boosting	
Intercept	25,3518	Intercept	-22,6566	Intercept	-32,6159	Intercept	-2,6410	avg_air_temp_17_18	
ndvi_17_18	89,6458	ndvi_17_18	89,3246	ndvi_17_18	89,3246	ndvi_17_18	74,0670	avg_air_temp_21_22	
ndvi_19_20	-29,0437	ndvi_19_20	-28,5187	ndvi_19_20	-28,5187	ndvi_27_28	10,4961	ndvi_17_18	
ndvi_23_24	74,2298	ndvi_21_22	3,3613	ndvi_21_22	3,3613	precip_17_18	-7,3233	ndvi_19_20	
ndvi_25_26	-89,5784	ndvi_23_24	70,5142	ndvi_23_24	70,5142	precip_19_20	-0,2855	ndvi_21_22	
ndvi_27_28	52,0282	ndvi_25_26	-89,1795	ndvi_25_26	-89,1795	temp_21_22	-0,9322	ndvi_23_24	
temp_17_18	0,3297	ndvi_27_28	51,8736	ndvi_27_28	51,8736	RASE Train, %	19,4	ndvi_25_26	
temp_19_20	0,7105	precip_17_18	-2,8866	precip_17_18	-11,1327	RASE Valid, %	22,1	ndvi_27_28	
temp_21_22	-4,1177	precip_19_20	-0,5677	precip_19_20	-1,4621	MAX E Train	15,7	RASE Train, %	12,7
RASE Train, %	18,2	precip_21_22	1,2401	temp_21_22	0,5720	MAX E Valid	12,6	RASE Valid, %	17,3
RASE Valid, %	19,4	temp_21_22	-0,7353	RASE Train, %	18,2			MAX E Train	10,6
MAX E Train	13,6	temp_23_24	1,3401	RASE Valid, %	19,5			MAX E Valid	13,0
MAX E Valid	11,7	temp_25_26	-0,7023	MAX E Train	13,6				
		RASE Train, %	18,2	MAX E Valid	11,7				
Regression Decision Tree		RASE Valid, %	19,5						
Intercept	41,6481	MAX E Train	13,6						
ndvi_17_18	94,5551	MAX E Valid	11,7						
temp_21_22	-3,9546								
RASE Train, %	19,4								
RASE Valid, %	22,5								
MAX E Train	14,3								
MAX E Valid	13,2								

Рис. 6.20. Опис моделей залежності урожайності озимої пшениці від
NDVI, погодних факторів

Для дослідження урожайності озимої пшениці під впливом факторів зовнішнього середовища було побудовано дві мережі Байєса. Для побудови першого набору мереж Байєса вихідні дані були переведені в дискретну форму за методом Optimal Binning (дискретизація проводилась за допомогою аналітичної системи SAS Enterprise Miner [8]).

Для побудови другої мережі Байєса, робились такі перетворення:

1. Взяті перші різниці від даних показників NDVI, температури та кількості опадів.
2. Переведені в дискретну форму за методом рівних проміжків.

На основі дискретизованих даних, мережа Байєса будувалась за допомогою програми GeNIe, використовуючи метод Greedy Thick Thinning. На рисунку 5 зображено мережу Байєса для озимої пшениці. Точність класифікації становить 47,06%. За побудованою мережею врожайність безпосередньо залежить від показників: NDVI на 17-18 тижні; різниці NDVI на 25-26 і 23-24 тижні; різниці NDVI на 27-28 і 25-26 тижні; різниці кількості опадів на 19-20 і 17-18 тижні; різниці температури на 21-22 і 19-20 тижні.

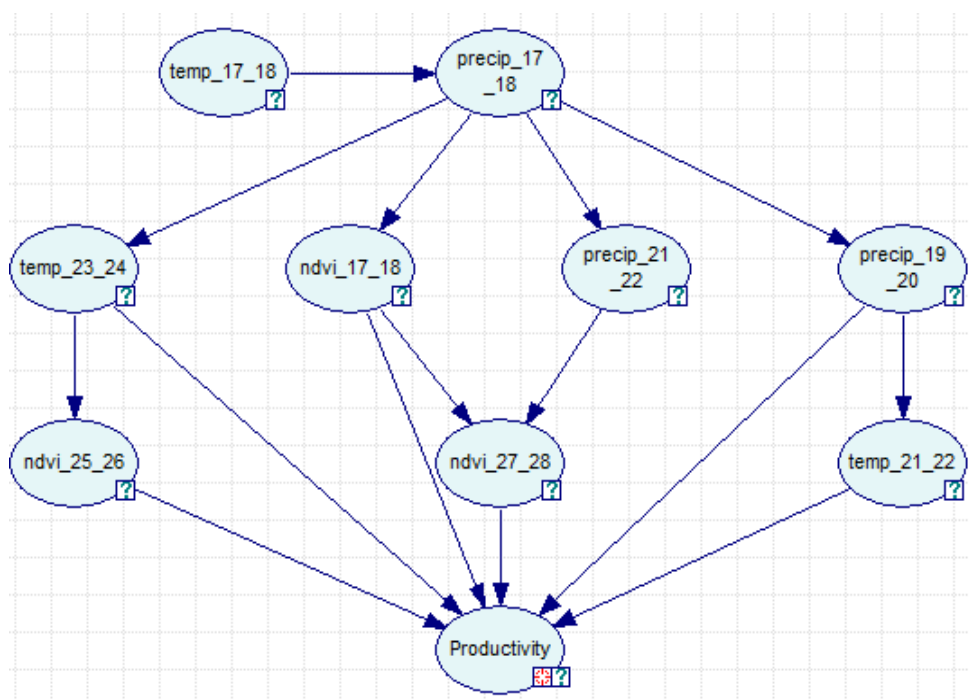


Рис. 6.21. Мережа Байєса (озима пшениця, взяті перші різниці вхідних даних)

У таблиці 8 представлено зведену інформацію про показники якості усіх побудованих моделей залежності врожайності озимої пшениці від NDVI і погодних факторів.

Таблиця 6.18

Значення показників якості регресійних моделей залежності врожайності озимої пшениці від NDVI і погодних факторів

Вибірка	Навчальна			Перевірочна		
Модель	RASE	SSE	MAX E	RASE	SSE	MAX E
Regression	5,21	977,74	13,64	5,54	459,81	11,73
LARS	5,21	977,73	13,64	5,57	465,68	11,74
LARS за критерієм мінімального SBC	5,54	1104,23	15,74	6,3	595,91	12,65
PLS	5,21	977,73	13,64	5,57	465,68	11,74
Gradient Boosting	3,63	475,25	10,58	4,94	366,78	13,04

Як видно з таблиці 6.18, найкращою виявилась модель, побудована за методом градієнтного бустингу. Для порівняння якості побудованих регресійних моделей з мережами Байєса, розраховано такі показники:

- для мереж Байєса розраховувалось відношення правильно класифікованих значень врожайності і вибиралась та мережа Байєса, у якої це показник вище;
- для регресійних моделей, вхідні дані показника врожайності і показника врожайності, розрахованого за відповідною моделлю, трансформувались в інтервали відповідно до мережі Байєса, з якою проводилось порівняння.

Таблиця 6.19

Порівняння якості побудови регресійних моделей і мереж Байєса (урожайність пшениці)

Модель	% класифікації на тестовому наборі даних
Мережа Байєса	78,43
Gradient Boosting	86,11
Regression	69,44
PLS	69,44
LARS	69,44

Досліджуючи отримані результати моделювання залежності врожайності озимої пшениці від NDVI та погодних факторів на основі регресійних моделей,

можна зробити висновок, що найкраща модель отримана за методом градієнтного бустингу (показники якості на навчальній вибірці такі: $RASE = 3,63$; $SSE = 475,25$; $MAX E = 10,58$; на перевіірочній вибірці: $RASE = 4,94$; $SSE = 366,78$; $MAX E = 13,04$), інші моделі показали гірші і схожі результати ($RASE$ від 5,21 до 5,54, SSE від 977,73 до 1105,32, $MAX E$ від 13,64 до 15,74; на перевіірочній вибірці: $RASE$ від 5,54 до 6,42, SSE від 459,81 до 595,91, $MAX E$ від 11,73 до 13,16).

Також побудовані мережі Байєса за двома методами – за вихідними даними, які переведені в інтервальний вимір за методом оптимального розбиття на групи на основі алгоритму дерева рішень. Для побудови мереж використовувався спеціальний компонент Optimal Binning в системі SAS Enterprise Miner [8] за вихідними даними, які переведені в різницеву форму і та в інтервальний вимір за методом рівних інтервалів. Отримані мережі Байєса порівнювались за загальною точністю класифікації, показники якої склали 55,36% та 70,98% для кукурудзи, і 78,43% та 47,06% для пшениці (відповідно до моделі). Порівняно результати, отримані за регресійними моделями та байєсівськими мережами: найкращі результати залежності врожайності для обох досліджуваних культур – ті, що отримані за методом градієнтного бустингу.

6.6 Дослідження ефективності управління розвитком локальних соціально-економічних систем – об'єднаних громад

В якості прикладу інформаційної технології для ідентифікації невизначеностей різних типів та їх розкриття у задачах підтримки прийняття рішень, розглянуто таку, що призначена для управління розвитком локальних соціально-економічних систем – об'єднаних громад, отримана з порталу даних Державний веб-портал бюджету для громадян (<https://openbudget.gov.ua>). Особливістю даної задачі є те, що процеси, які мають місце у таких системах, в більшості, є нелійнійними, а в процесі дослідження доводиться опрацьовувати невизначеності різних типів. Найбільш проблемними є питання опрацювання

пропусків даних та розв’язання задачі «прокляття розмірності». Адже часові ряди таких соціально-економічних показників об’єднаних громад містять пропуски, зазвичай є короткими, а досліджувану систему характеризує значна кількість показників.

Для прогнозного моделювання спроможності громад використано як статистичні дані щодо соціально-економічного розвитку громад, показники фінансової стійкості місцевих бюджетів, результати порівняльного аналізу ефективності бюджетних програм, так і розширений набір показників, що застосовуються у бюджетному процесі. У дослідженні використано дані більш, ніж 30 громад. Для відбору громад було виконано кластерний аналіз. Для дослідження відібрані громади, що увійшли в один кластер. Зв’язки між даними та приховані закономірності були розглянуті, використовуючи процедури інтелектуального аналізу даних. Крім того, на етапі збору та попередньої обробки даних, для кращого розуміння предметної області виділенні найбільш значимих для аналізу факторів було побудовано SWOT-таблицю. Використані в аналізі дані були попередньо підготовлені для подальшого використання для побудови моделей за допомогою процедури ETL. Для прогнозування показника “виконання бюджету за видатками” кращою виявилась модель лінійної регресії із вибором регресорів в автоматичному режимі за методом Stepwise (рис. 6.22).

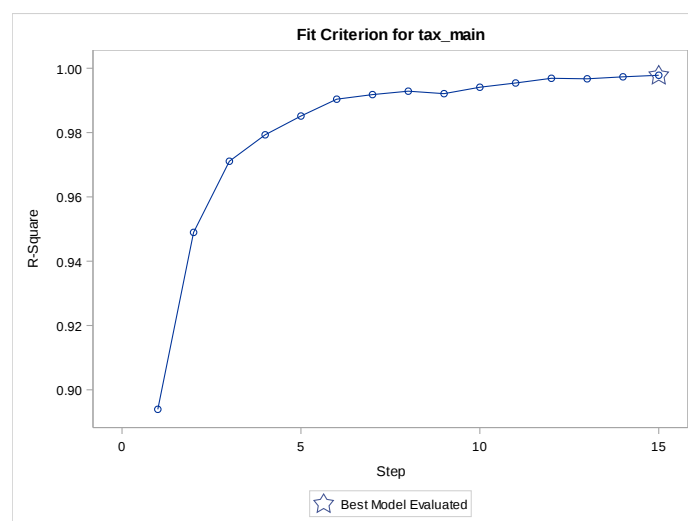


Рисунок 6.22 – Зміна статистики R^2 моделі в залежності від вибору регресорів, для прогнозування показника “виконання бюджету за видатками”.

Оцінка коефіцієнтів моделі для прогнозування показника “виконання

бюджету за видатками представлена в таблиці 6.20.

Пропуски, наявні в даних були заповнені, інтервальні змінні приведено до вигляду категоріальних, аномальні значення виключено з аналізу. Для усунення мультиколінеарності використано метод головних компонентів. Для заповнення пропусків даних використано методи знаходження середнього між першим попереднім та першим наступним значеннями ряду, моделі регресії (за малої, до 5% кількості пропусків) та значної (20%) кількості пропусків. Для оцінки якості регресійних моделей, застосованих для заповнення пропусків використовувався показник MAPE.

Таблиця 6.20 – Оцінка коефіцієнтів моделі для прогнозування показника “виконання бюджету за видатками”.

Назва показника (статті видатків)	Parameter Estimate	Standard Error	Type II S S	F-value	Pr > F
Вільний член	254.106	169.69	197288	2.24	0.1537
Реалізація загальнодержавні функції	0.99424	0.21527	1876760	21.33	0.0003
Житлово комунальні послуги	8.16851	0.49251	24201654	275.07	<.0001
Охорона здоров'я	3.59989	0.35915	8839193	100.47	<.0001
Фінансування власних повноважень	-0.39264	0.14507	644540	7.33	0.0156
Загальна середня освіта	0.23663	0.09350	563553	6.41	0.0222
Фінансування діяльності органи державного управління	1.61023	0.38901	1507459	17.13	0.0008
Комунальне господарство	-4.21820	0.64945	3711586	42.19	<.0001
Фінансування утримання інших закладів	1.18862	0.26600	1756854	19.97	0.0004
Культура і мистецтво	1.98848	0.51266	1323688	15.04	0.0013
Проведення виборів	9.71807	2.49986	1329609	15.11	0.0013
Інші видатки	0.85815	0.45286	315930	3.59	0.0763

Кінцевою метою використання пропонованої інформаційної технології було прогнозування доходної та витратної частини місцевого бюджету об'єднаних територіальних громад. Проаналізувати якість отриманого прогнозу можна за даними рис. 6.23.

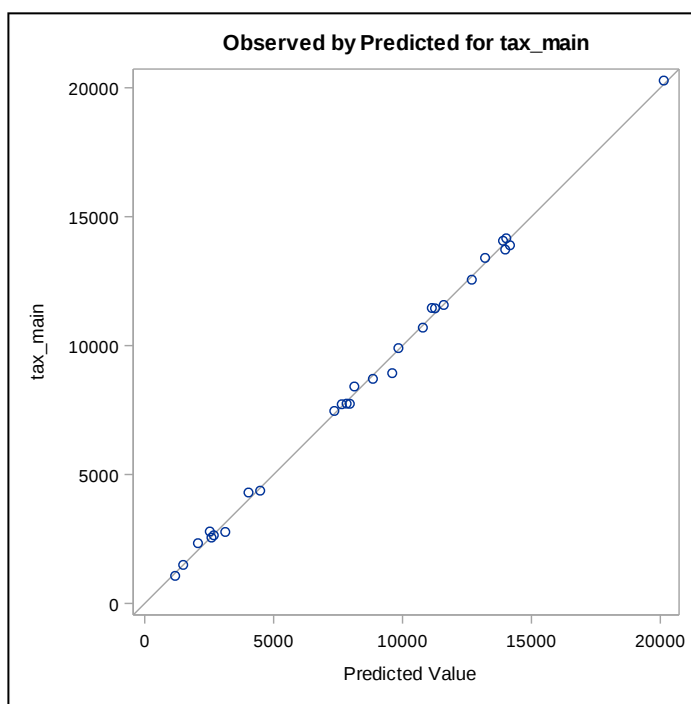


Рисунок 6.23 – Розподіл реальних та прогнозованих значень показника
«виконання бюджету за видатками»

Отримана математична модель має наступні прогнозуючі характеристики: $RMSE = 224$, $MAPE = 3,22\%$, $R^2 = 0,99$.

Для прогнозування показника “виконання бюджету за доходами” було побудовано модель лінійної регресії із вибором регресорів в автоматичному режимі за методом Stepwise (таблиця 6.21).

Таблиця 6.21 – Оцінка коефіцієнтів моделі для прогнозування показника “виконання бюджету за доходами”

Назва показника (статті видатків)	Parameter Estimate	Standard Error	Type II SS	F-value	Pr > F
Вільний член	590.70325	560.13206	2185481	1.11	0.3026
Податкові надходження	0.92724	0.11791	121536845	61.85	<.0001
Офіційні трансферти	1.82459	0.56346	20605728	10.49	0.0036
Субвенції з державного бюджету	-3.38282	0.77284	37649943	19.16	0.0002
Податок на майно	5.37428	0.84544	79408739	40.41	<.0001

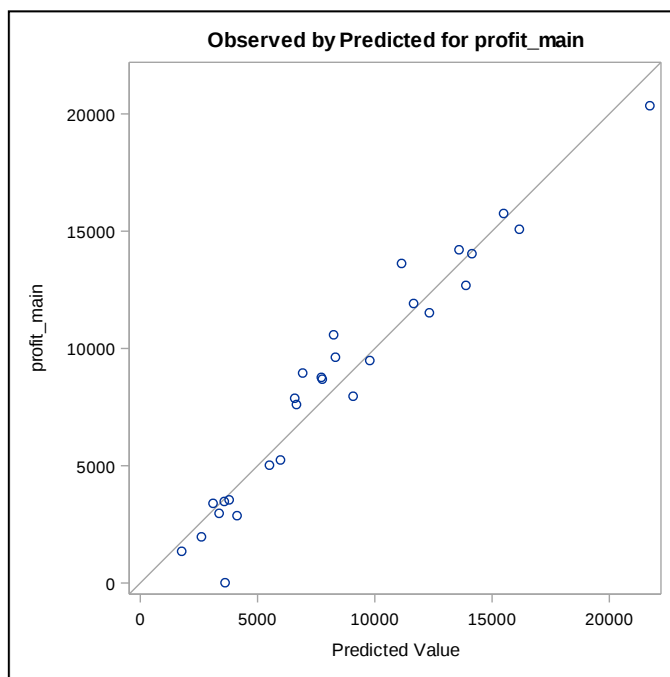


Рисунок 6.24 – Розподіл реальних та прогнозованих значень показника “виконання бюджету за видатками”.

Отримана математична модель має прийнятні прогнозуючі характеристики: $RMSE = 1074$, $MAPE = 12,7 \%$, $R^2 = 0,93$.

Таким чином, отримані результати дають можливість визначити проблеми та перспективи розвитку місцевого самоврядування: чи буде громада спроможною, чи зможе забезпечити соціальні гарантії мешканцям громади.

Для забезпечення зворотного зв'язку з жителями громади, в роботі запропоновано використання інтелектуального аналізу тексту – повідомлень жителів громади, розміщених у соціальних мережах, зверненнях, форумах, тощо. Приклад аналізу та категоризації у програмному середовищі SAS Text Miner.

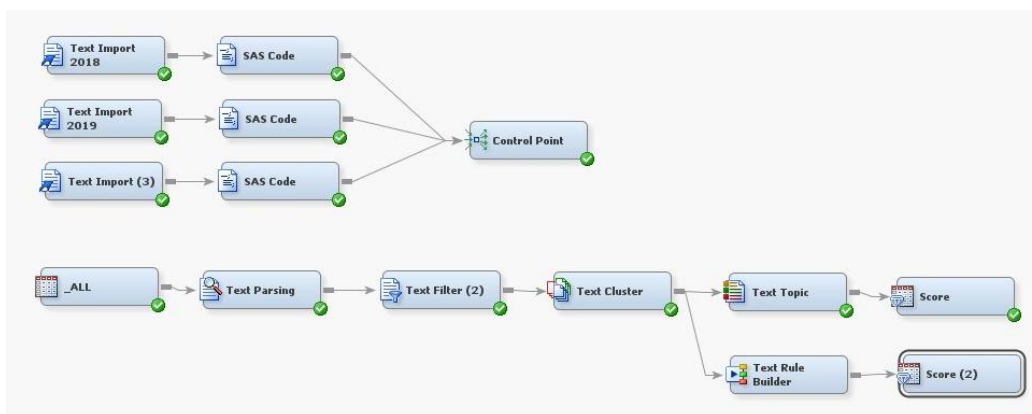


Рисунок 6.25 –Схема технологічного процесу аналізу текстових повідомлень у системі SAS Text Miner

Дослідження працездатності системи виконано на матеріалах груп повідомлень 20 територіальних громад (близько 10000 повідомлень). Для формування семантичного простору використані лексеми, сформовані з іменників та комбінацій іменників і дієслів. Шаблони словника, що містять терміни бюджетних програм та ієрархії предметної області склали основу тезаурусу, правила якого побудовані на основі виявлених в ході аналізу стоп-слів. Для кластеризації документів обрано алгоритм ієрархічної кластеризації (Hierarchical clustering), реалізований в модулі Text Mining SAS Enterprise Miner. В основі алгоритму використовується підхід на основі власних сингулярних чисел (SVD – singular value decomposition), що трансформуються у вагові коефіцієнти з частотної матриці термів корпусу документів.

Таблиця 6.22 - Статистичні характеристики автоматичної ієрархічної кластеризації документів

№ п/п	Режим ієрархічної кластеризації	Розмір SVD простору	Кількість отриманих кластерів	Час обчислення (секунд)	Якість кластеризації
1	low SVD resolution	35	10	53	прийнятна
2	medium SVD resolution	54	13	75	хороша
3	high SVD resolution	75	8	92	висока

Приклад формування концептуальних-зав'язків між термінами наведений на рис. 6.26.

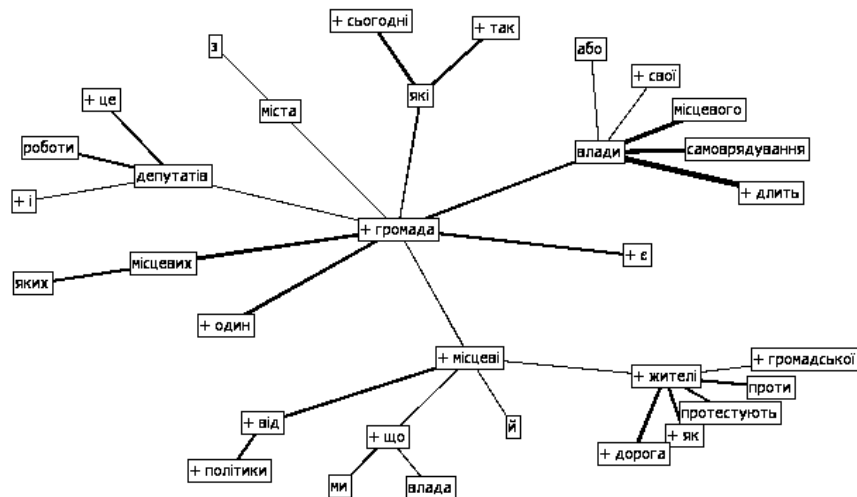


Рисунок 6.26 - Формування концептуальних-зав'язків для терміну «грумада», на основі корпусу з 10.000 інформаційних текстів.

Таблиця 6.23 – Результати ієрархічної кластеризації в режимі high SVD resolution.

Номер кластеру	Терми, що визнають кластер	Кількість документів	Процент документів
7	+держави будинків +бюджет ` city mandarin-news акцію аналізу багатоквартирних ми +більше які +можна +області +питання	2400	0,32
9	+результат +контроль +початок +рейтинг + аматорського +некомпетентна	1000	0,13(3)
10	+дослідження +діджиталізація +звітувати +змі	400	0,053(3)
12	+відповідно +реалізації +громадян +органів +загрозу +законом	1400	0,186(6)
16	+місцевих +народ +депутатів +країни +некомпетентна +виборців +голова	1000	0,13(3)
17	+асфальт +базар +бракує +демонструє +децентралізації	500	0,06(6)
18	+агломерації +будівництв +впритул +газозаправного +експерта +законопроект	400	0,053(3)
19	+місцеві +області +політики +послуг +аграрного +агронома +аспірантуру +багатоповерхівки +благоустрою	400	0,053(3)

Виявлено, що найбільш обговорюваними темами, що цікавлять мешканців є наявність транспортного сполучення (17,4%), проблеми з працевлаштуванням (35,8%), проблем роботи служб з надання адміністративних послуг (46,8%). На другому рівні альтернативи, пов'язані з створенням об'єднаних громад. Найбільш важливими альтернативами є забезпечення наповнюваності бюджетів - 38%, згорання медичної галузі - 24%, скорочення кількості закладів освіти – 27%, а інші - 11%.

Щодо емоційного забарвлення досліджуваної інформації, то найбільш негативні емоції викликала тематика "об'єднання громад" та "розширення міської забудови". Позитивні емоції – "пільгові перевезення" та "ремонт доріг". Це означає, що створення проблеми безробіття, погане транспортне сполучення та проблеми, пов'язані з роботою органів місцевого самоврядування та скороченням фінансування освіти та медицини з державного бюджету створюють пріоритет для формування об'єднаних громад. Виявлені інформаційні тенденції використовуються для виявлення наступного чинника: підвищеної тривоги та невизначеності населення в результативності реформ та забезпеченні соціальних гарантій. Це дає можливість сформулювати пропозиції щодо першочергових проектів, які необхідно включити до бюджету об'єднаних територіальних громад на наступний бюджетний рік.

Впровадження комп'ютерних інформаційних систем, оснований на створеному інструментарії, у роботу органів державного управління, місцевого самоврядування, дорадчих структур стане в нагоді при ідентифікації об'єктів критичної інфраструктури, коротко- та середньостроковому прогнозуванні, формуванні перспективних планів, отриманні стратегічних екологічних оцінок щодо природокористування при реалізації складних інфраструктурних проектів, транспортної інфраструктури, виробничої логістики, уточненні оцінок впливу на довкілля потенційно небезпечних для довкілля проектів, а також можуть бути використані і органами місцевого самоврядування для оцінювання тенденцій розвитку певних громад та оцінюванні результативності урядування.

Висновки до розділу 6

1. Подано представлено реалізацію методології проектування і впровадження інформаційної технології підтримки прийняття рішень в управлінні регіональним розвитком при розв'язанні основних задач регіонального управління. А саме:

- при розв'язанні задач коротко-, середньо- та довгострокового управління навантаженням енергосистеми при якому враховано цикли: годинні, добові, місячні та річні, а також, зміни енергоспоживання населенням та бізнесом у передсвяткові, святкові та післясвяткові дні, час доби, температуру навколишнього середовища, географічне розташування споживачів;
- при розв'язанні задач управління аграрною галуззю за врахуванням комбінованого впливу біологічних, екологічних, економічних та антропогенних складових;
- для управління розвитком локальних соціально-економічних систем – об'єднаних громад, за врахуванням нелінійності процесів і необхідністю опрацювання невизначеностей різних типів.

2. Запропоновані інструментальні засоби для розв'язання прикладних завдань прогнозування та підтримки прийняття рішень у різних галузях, створені на основі запропонованих методів, моделей і технологій аналізу даних та експертних оцінок, що розширить можливості практичного використання результатів дослідження.

3. Встановлено, що створена система може бути використана як окремо, так і доповнювати до існуючих системи прогнозування, застосовувані в державному управлінні та місцевому самоврядуванні.

ВИСНОВКИ

На основі виконаних теоретичних і експериментальних досліджень вирішено важливу науково-прикладну проблему в галузі інформаційних технологій – підвищення ефективності прийняття рішень в управлінні розвитком регіонів та громад в умовах невизначеності та ризику на основі релевантних методів і засобів аналізу даних, моделювання та прогнозування. При цьому зроблено наступні висновки та отримано такі результати:

1. Виконано системний аналіз сучасних методів моделювання та прогнозування процесів регіонального розвитку, які мають різну природу та протікають в умовах нестаціонарності та невизначеності, що дозволило визначити їхні обмеження та сформулювати вимоги до нових підходів.

2. Розроблено єдину формалізовану методологію побудови математичних моделей для опису та прогнозування соціально-економічних процесів регіонів і громад в умовах невизначеності, яка охоплює повний цикл від збору та інтеграції різнотипних даних до формування оптимальних управлінських рішень.

3. Удосконалено методи прогнозування соціально-еколого-економічних процесів різної природи на основі комбінованого використання ймовірнісних та регресійних моделей, що дозволило зменшити похибку прогнозування на 5–30 %.

4. Запропоновано метод оцінювання параметрів математичних моделей та їх ансамблів, який усуває зміщеність оцінок прогнозів і підвищує статистичні показники адекватності на 5–20 %.

5. Розроблено метод виявлення та розкриття невизначеностей різних типів за допомогою ймовірнісно-статистичних підходів, який забезпечує коректне прогнозування розвитку складних соціально-економічних процесів.

6. Побудовано ансамблі математичних моделей процесів, характерних для регіонів певного типу, та виконано їх застосування для прогнозування розвитку і підтримки прийняття рішень в умовах інформаційної невизначеності та ризику.

7. Розроблено багатомодельний і багатокритеріальний підхід до інтеграції різнотипної інформації, який базується на системному використанні методів аналізу, автоматизованого інтегрування, моделювання, прогнозування та багатокритеріального прийняття рішень з урахуванням динаміки процесів.

8. Виконано апробацію запропонованих підходів до моделювання та прогнозування, доведено їх здатність адекватно описувати причинно-наслідкові зв'язки між чинниками різних груп та підвищувати точність прогнозів на 10–20 %.

9. Проведено оцінювання адекватності побудованих моделей і сформульовано висновки щодо можливості їх використання для оцінювання та прогнозування розвитку соціально-еколого-економічних процесів як на рівні регіону, так і громади.

10. Розроблено методику розв'язання багатокритеріальних задач прийняття рішень із використанням процедур оптимізації, зокрема еволюційних та генетичних алгоритмів, інтегрованих у систему підтримки прийняття рішень.

11. Створено та реалізовано інформаційну технологію управління регіональним розвитком на основі запропонованих методів і моделей, придатну для впровадження у системи підтримки прийняття рішень органів державного управління та місцевого самоврядування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Trofymchuk O. M., Bidiuk P. I., Prosiiankina-Zharova T. I., Terentiev O. M. Decision support systems for modelling, forecasting and risk estimation: monography. Riga: LAP LAMBERT Academic Publishing. 2019. 176 p. ISBN: 978-3-659-61214-5
2. Trofymchuk O.M., Bidyuk P.I., Prosiiankina-Zharova T.I., Terentiev O.M. Bayesian data analysis in modeling and forecasting nonlinear nonstationary financial and economic processes. *Проблеми керування та інформатики*, 2023. № 4. С. 71-83. <https://doi.org/10.34229/1028-0979-2023-4-6>
3. Терент'єв О. М., Кириченко В. Е., Связінська Н. О., Просянкіна-Жарова Т. І. Прогнозування фінансових ризиків з використанням наївного і доповненого деревом класифікаторів на основі байєсівських мереж. *Наукові вісті НТУУ КПІ*. 2016. №2. С. 60-68. <https://doi.org/10.20535/1810-0546.2016.2.63882>
4. Терент'єв О. М., Просянкіна-Жарова Т. І., Саваст'янов В. В. Використання засобів текстової аналітики як інструменту оптимізації підтримки прийняття рішень у задачах розробки планів соціально-економічного розвитку України. *Реєстрація, зберігання і обробка даних*. 2016. Т. 18. № 3. С. 75-86. URL: http://nbuv.gov.ua/UJRN/rzod_2016_18_3_10
5. Терент'єв О. М., Просянкіна-Жарова Т. І., Саваст'янов В. В. Застосування когнітивного та ймовірнісного моделювання у задачах формування сценаріїв розвитку соціально-економічних систем. *Наукові вісті НТУУ КПІ*. 2016. №5. С. 37-47. <https://doi.org/10.20535/1810-0546.2016.5.79876>
6. Бідюк П. І., Терент'єв О. М., Просянкіна-Жарова Т. І., Ефендієв В. В. Прогнозне моделювання нелінійних нестационарних процесів у рослинництві з використанням інструментів SAS Enterprise Miner. *Наукові вісті НТУУ КПІ*. 2017. №1. С. 24-36. <https://doi.org/10.20535/1810-0546.2017.1.87423>
7. Bidyuk P., Overmyer S., Prosyankina-Zharova T., Terentiev O. Methodology of Modeling and Forecasting Nonlinear Processes in Finances. *Наукові вісті НТУУ*

- KIII*. 2018. №1. С. 15 – 25. <https://doi.org/10.20535/1810-0546.2018.1.120361>
8. Bidyuk P., Prosyankina-Zharova T., Terentiev O., Medvedeva M. Adaptive modelling for forecasting economic and financial risks under uncertainty in terms of the economic crisis and social threats. *Технологічний аудит та резерви виробництва*. 2018. №4. С. 24-36. <https://doi.org/10.15587/2312-8372.2018.135483>
 9. Bidyuk P. I., Korshevnyuk L. O., Gozhyi O. P., Kalinina I. O. , Prosyankina-Zharova T. I., Terentiev O. M. Modeling and forecasting financial and economic processes with decision support system. *Наукові вісми KIII*. 2019. № 5-6. С. 7-17. <https://doi.org/10.20535/kpi-sn.2019.5-6.176835>
 10. Bidyuk P.I., Prosyankina-Zharova T.I., Diakon V., Diakon D. Modeling and forecasting nonlinear nonstationary processes in financial sphere. *Technology Audit and Production Reserves*. 2023. Vol. 4. № 2(72), P. 37 – 46. <https://doi.org/10.15587/2706-5448.2023.286516>
 11. Terentiev, O., Prosyankina-Zharova, T., Diakon, V., Manuilenco, R. Development of mathematical models to support decision-making regarding the functioning of critical infrastructure in the industry of energy supply. *Technology Audit and Production Reserves*, 2023. 6(2(74), 44–49. <https://doi.org/10.15587/2706-5448.2023.293205>
 12. Лисов Б., Гуськова В., Просянкіна-Жарова Т., Халигов А. Інформаційна технологія класифікації даних моніторингу систем і мереж передачі даних для виявлення кіберзагроз. *Вісник Національного технічного університету "ХПІ". Серія: Інформатика і моделювання*. 2025. №.1. Вип. 1(13). С. 128-147. <https://doi.org/10.20998/2411-0558.2025.01.09>
 13. Просянкіна-Жарова Т. Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком. Екологічна безпека та природокористування, 2025. №54(2), С. 164–184. <https://doi.org/10.32347/2411-4049.2025.2.164-184>
 14. Prosyankina-Zharova T., Terentiev O., Bidyuk P., Makukha M. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and

- forecasting. *Journal of Mathematics and System Science*. 2016. №6. P. 112-122. <https://doi.org/10.17265/2159-5291/2016.03.003> (CIIA, Index of Copernicus)
15. Bidyuk P. Modelling Nonlinear Nonstationary Processes in Macroeconomy and Finances/ Bidyuk P., Prosyankina-Zharova T., Terentiev O. / *Advances in Computer Science for Engineering and Education. ICCSEEA 2018. Advances in Intelligent Systems and Computing* / Ed. Hu Z., Petoukhov S., Dychka I., He M. (eds). Vol. 754. Cham: Springer, 2019. P.735-745. https://doi.org/10.1007/978-3-319-91008-6_72 (**Scopus Q3**). ISSN 2194-5357
 16. Bidiuk P. I., Prosyankina-Zharova T. I., Terentieev O. M., Lakhno V. A., Zhmud O. V. Intellectual technologies and decision support systems for the control of the economic and financial processes. *Journal of Theoretical and Applied Information Technology*. 2019. Vol. 97. No. 1. P. 71-87. (**Scopus Q3**) URL: <https://www.jatit.org/volumes/Vol97No1/7Vol97No1.pdf>
 17. Terentiev O. M., Prosyankina-Zharova T. I., Lahno V. A., Usatiuk Y. V. The features of the predictive computing modeling power system load in terms of reforming energy market. *Journal of Theoretical and Applied Information Technology*. 2020. Vol. 98. No. 2. P. 163-182. (**Scopus Q4**). URL: <https://www.jatit.org/volumes/Vol98No2/1Vol98No2.pdf>
 18. Terentiev O., Prosyankina-Zharova T., Savastiyanov V., Lakhno V., Kolmakova V. The features of building a portfolio of trading strategies using the SAS OPTMODEL procedure. *Computation*. 2021. 9(7). 77. <https://doi.org/10.3390/computation9070077> (**Scopus Q2**)
 19. Sharova, T., Zemlianska, A., Sharov, S., Prosyankina-Zharova, T., Kavun, L. Independent Evaluation of Learning Outcomes as a Tool of the Internal Quality Assurance System. *TEM Journal*. 2023, 12(2), P. 700–709. <https://doi.org/10.18421/TEM122-13> (**Scopus Q3**)
 20. Sharov S., Chakraborty S., Zemlianskyi A., Kolmakova V., Prosyankina-Zharova T. Quantitative Analysis of Online Courses for Learning the JavaScript Programming Language. *International Journal on Advanced Science, Engineering*

- and Information Technology*. 2024. Vol. 14, No. 4, pp. 1216-1225.
<https://doi.org/10.18517/ijaseit.14.4.19976> (Scopus Q4)
21. Prosyankina-Zharova T. I., Terentiev O. M., Bidyuk P. I., Gasanov A. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and forecasting: *On control and optimization with industrial applications (COIA-2015)*, Proceedings of the fifth international conference (Baku, 27-29 August 2015), Baku: Institute of applied mathematics BSU, 2015. P. 365-368. URL: http://coia-conf.org/upload/editor/files/COIA_2015.pdf
 22. Терент'єв О.М., Св'язінська Н.О., Прос'янкіна-Жарова Т.І. Візуалізація типології регіонів України із використанням система SAS Enterprise Guide 7.1. *Геоінформаційні системи і комп'ютерні технології еколого-економічного моніторингу: матеріали міжнар. наук.-техн. конф. (Дніпропетровськ, 13–15 квітня 2016 р. Дніпропетровськ: ГВУЗ «НГУ» МОН України, 2016. С. 21-25. URL: <https://ela.kpi.ua/handle/123456789/19565>*
 23. Text mining analysis of agriculture internet sources using SAS software / Prosyankina-Zharova T. I., Terentiev O. M., Bidyuk P. I., Sviazinska N. O., Kyrychenko V. E. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2016): матеріали міжнар. конф. (м. Залізний порт, 24-28 трав. 2016 р.). Херсон: ХНТУ, 2016. С. 244-246. URL: <https://ela.kpi.ua/handle/123456789/19575>*
 24. Ivanova Y. V., Terentiev O. N., Korshevnyuk L. O., Prosyankina-Zharova T. I. Using modified logistic regression to increase customer response rate to marketing campaigns. *International conference on System Analysis and Information Technology (SAIT 2016)*. Proceedings of the 18-th conference (Kyiv, May 30 – June 2). Kyiv: Institute for Applied System Analysis, National Technical University of Ukraine “KPI”, 2016. С. 304-305. URL: http://sait.kpi.ua/media/filer_public/73/32/7332a68e-e93b-4c57-a3c8-66f11ee074cd/sait2016ebook.pdf
 25. Terentiev O. M., Prosyankina-Zharova T. I. Development of scoring systems for decision-making in crediting individuals and legal entities. *Розвиток*

- підприємницької діяльності інтеграції України в світове господарство та інформатизації основних напрямків поступу.* Умань : Видавець “Сочінський М.М.”, 2017. С. 184-187. URL: <https://ela.kpi.ua/handle/123456789/19626>.
26. Терентьев О. М., Просьянкина-Жарова Т. І. Застосування прогностного моделювання для дослідження нелінійних нестационарних процесів у рослинництві. *Розвиток підприємницької діяльності інтеграції України в світове господарство та інформатизації основних напрямків поступу.* Умань : Видавець “Сочінський М.М.”, 2017. С. 113-116. URL: <https://ela.kpi.ua/handle/123456789/19619>.
27. Бідюк П. І., Терентьев О. М., Просьянкина-Жарова Т. І. Методи заповнення пропусків даних у задачах прогностного моделювання соціально-економічних процесів. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2017)* : матеріали міжнар. конф. (м. Залізний порт, 22-26 трав. 2017 р.). Херсон : ПП Вишемирський В. С., 2017. С. 185-187. URL: <https://ela.kpi.ua/handle/123456789/19679>.
28. Бідюк П. І. Терентьев О. М., Просьянкина-Жарова Т. І., Савастьянов В. В. Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни. *International conference on System Analysis and Information Technology (SAIT 2017). Proceedings of the 19-th. conference (Kyiv, May 22 – 25. 2017)..* Kyiv: National Technical University of Ukraine “KPI”, 2017. С. 253-255. URL: http://sait.kpi.ua/media/filer_public/38/b9/38b90ff9-322e-4190-b445-e58ff05759fc/sait2017ebook.pdf
29. Bidyuk P. I., Prosyankina-Zharova T.I., Terentiev O.M. Multi model adaptive forecasting using decision support system *Моделювання та прогнозування економічних процесів.* Матеріали XI Науково-практичної конференції.(м. Київ, грудень 6-8 грудня 2017 р.). К.: КПІ ім. Ігоря Сікорського, 2017, С. 14-17. URL <https://ecocyber.fmm.kpi.ua/wp-content/uploads/2022/06/mpep2017.pdf>

30. Трофимчук О.М., Бідюк П. І., Просянкін-Жарова Т. І., Терент'єв О. М. Адаптивне моделювання і прогнозування у системах підтримки прийняття рішень сільськогосподарського призначення. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях*: колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2017. С. 21-28. URL: <https://itgip.org/wp-content/uploads/2018/08/zbirka-maket-1.pdf>
31. Bidyuk, P. Terentiev O., Prosyankina-Zharova T. Dynamic processes forecasting and risk estimation under uncertainty using decision support systems. IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Proceedings of the Proceedings of the conference (Kyiv, 29 May-2 June 2017), Kyiv: Igor Sikorsky Kyiv Polytechnic Institute. 2017. P. 795-800. (Scopus) <https://doi.10.1109/UKRCON.2017.8100355>. ISBN:978-1-5090-3006-4
32. Trofymchuk O. M. Forecasting nonstationary processes in demography, ecology, economy and finances/O. M. Trofymchuk, P. I. Bidyuk, T. I. Prosyankina-Zharova, O. M. Terentiev *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях*: колективна монографія/ ред. С.О. Довгий]. К.: ТОВ “Видавництво Юстон”, 2018. С. 113-123. URL: <https://itgip.org/wp-content/uploads/2018/10/zbirka-maket-2018-1.pdf>
33. Трофимчук О.М., Бідюк П.І., Терент'єв О.М., Просянкін-Жарова Т.І. Особливості прогнозного моделювання з можливістю картографування та інтелектуального аналізу геопросторових даних у еколого-економічних дослідженнях; XVIII-а Міжнародна науково-практична конференція: «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: актуальні питання», колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2019. С. 18-26. URL: https://itgip.org/wp-content/uploads/2019/10/1_knyga_sajt.pdf

- 34.Шолохов О. В., Терентьев О. М., Просьянкіна-Жарова Т. І. Підвищення ефективності соціальних комунікацій на основі аналізу інтернет-джерел засобами text mining. *Прикладні системи та технології в інформаційному суспільстві: матеріали міжнар. наук.-практ. конф.* (Київ, 30 вер. 2020 р.). К.: Київський нац. ун-т імені Тараса Шевченка, 2020. С. 241-245. URL: https://aistis.knu.ua/wp-content/uploads/2020/11/tezi_2020_.pdf
- 35.Trofymchuk O.M., Bidyuk P.I., Prosyankina-Zharova T.I., Terentiev O.M. Decision support system for modeling and forecasting ecological processes. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: колективна монографія* / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2020. С. 23-44. URL: https://itgip.org/wp-content/uploads/2020/10/zbirka_2020_1.pdf
- 36.Korbicz J., Bidyuk P., Kuznietsova N., Terentiev O., Prosiankina-Zharova T. Multivariate distribution model for financial risks management. *The 9th International Conference "Information Control Systems & Technologies"(ICST2020)* (Odessa, Sept. 24–26, 2020). CEUR Workshop Proceedings, 2020. P. 416-429. URL: <https://ceur-ws.org/Vol-2711/paper32.pdf> (Scopus).
- 37.Trofimchuk O., Prosiankina-Zharova T.; Bidiuk P., Terentiev O. Operational risk estimation using probabilistic model. *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)* (Tashkent, Oct. 07-09, 2020), P. 1-5. <https://doi.10.1109/AICT50176.2020.9368630> (Scopus) ISBN:978-1-7281-7386-3, ISSN:2472-8586. URL: <https://ieeexplore.ieee.org/document/9368630>
- 38.Korbicza J., Trofymchuk O., Bidiuk P., Kuznietsova N., Terentiev O., Prosiankina-Zharova T. Probabilistic Modeling of Risks of Different Origin. *Modern Information Technologies for Environmental Safety Management, Nature Management, Emergency Measures*, October 7, 2021, Kyiv, Ukraine 2021. P. 29-42. URL: <http://ceur-ws.org/Vol-3021> (Scopus)

39. Трофимчук О. М., Бідюк П. І., Терент'єв О. М., Просянкін-Жарова Т.І. Застосування методики аналізу подібності часових рядів для прогнозування процесів різних типів в умовах інформаційної невизначеності; XX-а Міжнародна науково-практична конференція: *«Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: виклики 2021 року»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2021. С. 19-25. URL: https://itgip.org/wp-content/uploads/2021/10/1_zbirka_2021.pdf
40. Терент'єв О. М., Просянкін-Жарова Т.І., Дякон Д. В. Застосування засобів опрацювання неструктурованих даних у задачах прогнозного моделювання; XX-а Міжнародна науково-практична конференція: *«Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: виклики 2021 року»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2021. С. 161-168. URL: https://itgip.org/wp-content/uploads/2021/10/1_zbirka_2021.pdf
41. Терент'єв О.М. Просянкін-Жарова Т.І. Шолохов О.В. Особливості застосування інструментів інтелектуального аналізу даних компанії SAS у торгівлі криптовалютою *Прикладні системи та технології в інформаційному суспільстві: матеріали міжнар. наук.-практ. конф. (Київ, 30 вер. 2022 р.)*. К.: Київський нац. ун-т імені Тараса Шевченка, 2022. С. 200-210. URL: <https://aistis.knu.ua/wp-content/uploads/2022/10/AISTIS-2022.pdf>
42. Trofymchuk O. M., Bidyuk P.I., Prosiankina-Zharova T. I., Terentiev O. M. Bayesian data analysis in modeling and forecasting nonlinear nonstationary financial and economic processes; XXI-а Міжнародна науково-практична конференція: *«Інформаційно-комунікаційні технології та сталий розвиток»*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2022. URL: https://itgip.org/wp-content/uploads/2022/11/tezy_6.docx
43. Дякон Д.В., Просянкін-Жарова Т.І. Проблеми застосування інформаційних технологій у міжвідомчій співпраці в сфері реагування на різні загрози й

- надзвичайні ситуації. *Інформаційно-комунікаційні технології та сталий розвиток*, колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2022. URL: https://itgip.org/wp-content/uploads/2022/11/tezy_32a.pdf
44. Терентьев О. М., Просьянкіна-Жарова Т. І. Опрацювання інформаційної невизначеності із застосуванням методики виявлення подібності процесів. *Виклики і загрози для критичної інфраструктури*. матеріали міжнар. Наук.-практ. конф. (м. Київ, 29-30 червня 2023 р.) URL: <https://conference.cyberspace.org.ua>
45. Халігов А.А., Просьянкіна-Жарова Т.І., Гуськова В. Г. Удосконалення інформаційної технології аналізу даних на порталі open budget *Інформаційно-комунікаційні технології для перемоги та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2023.. С. 73–76. URL: https://itgip.org/wp-content/uploads/2023/11/1_zbirka_08_11_23-1-1.pdf
46. Бідюк П. І., Терентьев О. М., Просьянкіна-Жарова Т.І. Інформаційна технологія побудови прогнозів функціонування об’єктів критичної інфраструктури в умовах кризової ситуації. *Інформаційно-комунікаційні технології для перемоги та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2023. С. 114–117. URL: https://itgip.org/wp-content/uploads/2023/11/1_zbirka_08_11_23-1-1.pdf
47. Вишняков В.Ю., Глуган Ф.В., Просьянкіна-Жарова Т. І. Інформаційна технологія визначення тематичних точок на підстильній поверхні Землі із використанням даних дистанційного зондування Землі, отриманих з різних джерел. *Математичне моделювання та інформаційно-комунікаційні технології для зміцнення та відновлення*. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2024. С. 126-129. https://itgip.org/wp-content/uploads/2024/11/2024-11-24_zbirka_all_07_11_2024_148x210.pdf
48. Халигов А.А., Просьянкіна-Жарова Т. І., Гуськова В. Г. Використання рекомендаційної системи для підтримки прийняття рішень у системі кібербезпеки об’єкта критичної інфраструктури *Математичне моделювання*

- та інформаційно-комунікаційні технології для зміцнення та відновлення. колективна монографія / ред. С.О. Довгий. К.: ТОВ “Видавництво Юстон”, 2024. С. 149-154. https://itgip.org/wp-content/uploads/2024/11/2024-11-24_zbirka_all_07_11_2024_148x210.pdf
- 49.Згуровський М. З., Бідюк П. І., Терентьев О.М., Просянкін-Жарова Т. І. Байєсівські мережі в системах підтримки прийняття рішень: навч. посіб. К: ТОВ «Видавниче підприємство «Едельвейс», 2015. 300 с. ISBN 978-966-2748-73-4. URL: <https://ela.kpi.ua/handle/123456789/19582>
- 50.Bates D. M., Watts D. G. Nonlinear Regression Analysis and Its Applications. New York: John Wiley & Sons, 1988. 392 p.
- 51.Arora S., Little M. A., McSharry P. E. Nonlinear and nonparametric modeling approaches for probabilistic forecasting of the US gross national product. *Studies in Nonlinear Dynamics and Econometrics*. 2013. Vol. 17, No. 4. P. 395–420.
- 52.Pearson R. K., Ogunaike B. A. Nonlinear process identification. Chapter 2. *Nonlinear process control*. Prentice-Hall, Inc., NJ, USA, 1997. P. 11–110. URL: <https://dl.acm.org/doi/book/10.5555/248020> (Last accessed: 18.07.2020).
- 53.Márcio G., Garcia P., Marcelo C., Medeiros M. C., Gabriel F. R. Vasconcelos. Real-time inflation forecasting with high-dimensional models: the case of Brazil. *International Journal of Forecasting*. 2017. Vol. 33, issue 3. P. 679–693.
- 54.Szeles M. R., Muñoz R. M. Analyzing the regional economic convergence in Ecuador. Insights from parametric and nonparametric models. *Romanian Journal of Economic Forecasting*. 2016, Vol. XIX (2). P. 43–65.
- 55.Zar J. H. Biostatistical Analysis. Third Edition. New York : Prentice-Hall, Inc., 1996. 662 p.
- 56.Demirtas H., Sally A., Recai M. Plausibility of multivariate normality assumption when multiply imputing non-gaussian continuous outcomes: a simulation assessment. Technical report 2006-001. Chicago, C: Division of Epidemiology and Biostatistics University of Illinois at Chicago., 2006. 23 p. URL: <http://dx.doi.org/10.1080/10629360600903866>.

57. Hoaglin D.C., Mosteller F., Tukey J.W. Understanding Robust and Exploratory Data Analysis. New York : John Wiley & Sons, 2000. 472 p.
58. John P. W. M. Statistical Design and Analysis of Experiments. New York : The Macmillan Company, 1971. 756 p.
59. Myers R. H. Response Surface Methodology. Virginia : Polytechnic and State University, 1988. 705 p.
60. Wilcox R.R. Keselman H.J. Modern robust data analysis methods: Measures of central tendency. *Psychological Methods*. 2003. Vol. 8 (3). P. 254-274.
61. Zgurovsky M. Z., Zaychenko Y. P. Big Data: Conceptual Analysis and Applications. International Publishing Switzerland, 2019. 277 p.
62. Zgurovsky M. Z., Zaychenko Y. P. The Fundamentals of Computational Intelligence: System Approach. Switzerland : Springer International Publishing, 2016. 375 p.
63. Zhang Zhengyou. Parameter Estimation Techniques: A Tutorial with Application to Conic. *INRIA Research Report*. 2004. №.2676, 34 p. URL: <http://research.microsoft.com/en-us/um/people/zhang/Papers/ZhangIVC-97-01.pdf> (Last accessed: 19.07.2020).
64. Азарсков В. Н., Житецкий Л. С., Николаенко С. А. О сходимости стандартных алгоритмов последовательного обучения нейросетевых моделей в нестохастической среде. *Нейроинформатика-2011* : сб. науч. трудов XIII Всерос. науч.-тех. конф. Москва : НИЯУ МИФИ, 2011. Ч. 2. С. 258–260.
65. Berry M. J. A., Linoff G. S. Data Mining Techniques. 2nd edition. New York : Wiley Publishing, Inc., 2004. 670 p.
66. Tangirala A. K. Non-stationary processes : Course of lectures “Introducing to random process”. IIT, Madras, India, 2021. 18 p.
67. Márcio G., Garcia P., Marcelo C., Medeiros M. C., Gabriel F. R. Vasconcelos. Real-time inflation forecasting with high-dimensional models: the case of Brazil. *International Journal of Forecasting*

68. Bidyuk P., Prosyankina-Zharova T., Terentiev O. Modelling Nonlinear Nonstationary Processes in Macroeconomy and Finances. *Advances in Computer Science for Engineering and Education. ICCSEEA 2018. Advances in Intelligent Systems and Computing* / Ed. Hu Z., Petoukhov S., Dychka I., He M. Cham : Springer, 2019. Vol. 754. P. 735–745. URL: http://doi.org/10.1007/978-3-319-91008-6_72
69. Коршевнік Л. О., Терент'єв О. М., Бідюк П. І. Методика побудови математичних моделей динамічних процесів. *Системний аналіз та інформаційні технології (SAIT 2013)* : матеріали 15-ї міжнар. наук.-техн. конф. (Київ, 27–31 трав. 2013 р.). Київ : ННК “ІПСА” НТУУ “КПІ”, 2013. С. 288–289.
70. Ивахненко А. Г., Юрачковский Ю. П. Моделирование сложных систем по экспериментальным данным. Киев : Радио и связь, 1987. 120 с.
71. Кашкарьов А. О., Терент'єв О. М. Аналіз випадкових процесів за допомогою швидкого перетворення Фур'є. *Праці Таврійського державного агротехнологічного університету*, 2010. Вип. 10, Т. 10. С. 158–162.
72. Бард Й. Нелинейное оценивание параметров : пер. с англ. Москва : Статистика, 1979. 349 с.
73. Sel'kov E. E. (1968). *Self-Oscillations in Glycolysis 1. A Simple Kinetic Model. European Journal of Biochemistry*. 1968. 4 (1). S. 79–86. URL: <https://febs.onlinelibrary.wiley.com/doi/full/10.1111/j.1432-1033.1968.tb00175.x> (Last accessed: 14.12.2020).
74. Polozhaenko S. A., Abdullach O. M. Methods of mathematical design and machine authentication of non – statics anomalous diffusive processes. *Informatics and mathematical method sinmodelling*. 2015. Vol. 5, N 3. P. 249–258.
75. Krak I., Barmak O., Manziuk E. Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology, *Computation alIntelligence*. 2020. Vol. 1. P. 1–26.

76. Roberts F. S., Discrete mathematical models with applications to social, biological and environmental objectives. *New York* : Englewood Cliffs : Prentice-Hall. 1986. XVI, 559 p.
77. Jordanova T. An Introduction to Stationary and Non-Stationary Processes. 2020. URL: <https://www.investopedia.com/articles/trading/07/stationary.asp#:~:text=A%20non%2Dstationary%20process%20with,constant%20and%20independent%20of%20time.&text=It%20specifies%20the%20value%20at,trend%2C%20and%20a%20stochastic%20component> (Last accessed: 21.06.2020).
78. Cheng C., Sa-Ngasoongsong A, Faruk O., Le T. Time Series Forecasting for Nonlinear and Nonstationary Processes: A Review and Comparative Study. *IIE Transactions*, 2015. DOI: 10.1080/0740817X.2014.999180.
79. Explained: Linear and nonlinear systems. *MIT News*. URL: <https://news.mit.edu/2010/explained-linear-0226> (Last accessed: 10.01.2021).
80. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. Москва : Финансы и статистика, 2003. 416 с.
81. Данилов В. Я, Бідюк П. І, Жиров О. Л., Бідюк О. П. Прогнозування кредитоспроможності фізичних осіб за допомогою байєсівських мереж. *Економіка: теорія та практика*. 2016. №2 (8). С. 56–65.
82. Литвиненко В. І., Кожухівська О. А., Фефелов А. О. Метод прогнозування гетероскедастичних процесів з використанням синтезованих поліноміальних нейронних мереж. *Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі*. 2015. №829. С. 201–214. URL: http://nbuv.gov.ua/UJRN/VNULPICM_2015_829_16 (дата звернення: 22.12.2020)
83. Терентьев О. М., Кириченко В. Е., Связінська Н. О., Просянкін-Жарова Т. І. Прогнозування фінансових ризиків з використанням наївного і доповненого деревом класифікаторів на основі байєсівських мереж. *Наукові вісті НТУУ КПІ*. 2016. №2. С. 60–68. URL: <https://doi.org/10.20535/1810-0546.2016.2.63882>

- 84.Снитюк В. Є. Прогнозування. Моделі. Методи. Алгоритми : навч. посіб. Київ : Маклаут, 2008. 364 с.
- 85.Светуньков И.С., Светуньков С.Г. Методы социально-экономического прогнозирования. Теория и методология. Москва: Юрайт, 2015. 351 с.
- 86.Brocklebank J. C., Dickey D. A. *SAS System for Forecasting Time Series / Second Edition* Cary N C. SAS Institute Inc., 2003. 418 p.
- 87.Bidyuk, P. Terentiev O., Prosyankina-Zharova T. Dynamic processes forecasting and risk estimation under uncertainty using decision support systems. IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON) : Proceedings of the Proceedings of the conference (Kyiv, 29 May-2 June 2017). Kyiv : Igor Sikorsky Kyiv Polytechnic Institute. 2017. P. 795–800. (Scopus). URL: <https://doi.8100355>
- 88.Badach A. Adaptive Models in Econometric Forecasting. IFAC Proceedings Volumes. 1980. Vol. 13, issue 5. P. 271–277.
- 89.Littell R. C., Milliken G. A., Stroup W. W., Wolfinger R. D. *SAS® System for Mixed Models*, Cary, NC : SAS Institute Inc., 1996. 828 p.
- 90.Dobrescu E. Modelling an Emergent Economy and Parameter Instability Problem. *Journal for Economic Forecasting of Institute for Economic Forecasting*. 2017. Vol. 0(2). P. 5–28.
- 91.Bidyuk P., Huskova V., Terentiev O. Client solvency estimation using intellectual data analysis approach, *Advanced Information Systems and Technologies (AIST 2018)* : Proceedings of the 6th International Conference (Sumy, 16–18 May 2018), Sumy : Sumy State University, 2018 . С. 12–16.
- 92.Bidyuk P., Prosyankina-Zharova T., Terentiev O., Medvedeva M. Adaptive modelling for forecasting economic and financial risks under uncertainty in terms of the economic crisis and social threats. *Технологічний аудит та резерви виробництва*. 2018. №4. С. 24–36. URL: <https://doi.org/10.15587/2312-8372.2018.135483>.
- 93.Bauer D. J. A Semiparametric Approach to Modeling Nonlinear Relations Among Latent Variables. *Structural Equation Modeling: A Multidisciplinary*

- Journal*. 2005. Vol. 12, issue 4. URL: https://www.tandfonline.com/doi/abs/10.1207/s15328007sem1204_1 (Last accessed: 11.08.2020).
94. Andreica, M., Popescu, M.M.E., Micu D., Albu E. Adaptive management procedural model for support of economic organizations. *Proc of 10th International Management Conference "Challenges of Modern Management*. 2016. P. 295–301.
 95. Bidiuk P. I., Prosiankina-Zharova T. I., Terentieev O. M., Lakhno V. A., Zhmud O. V. Intellectual technologies and decision support systems for the control of the economic and financial processes. *Journal of Theoretical and Applied Information Technology*. 2019. Vol. 96. No. 1. P. 71–87. (Scopus Q3)
 96. Bidiuk P. I., Menyailenko O. S., Polovtcev O. V. *Methods of Forecasting*, Lugansk : Alma Mater, 2008. 608 p.
 97. Bidiuk P., Overmyer S., Prosyankina-Zharova T., Terentieev O. Methodology of Modeling and Forecasting Nonlinear Processes in Finances. *Наукові вісники НТУУ КНІІ*. 2018. №1. С. 15–25. URL: <https://doi.org/10.20535/1810-0546.2018.1.120361>.
 98. Trofymchuk O. M., Bidiuk P. I., Prosiankina-Zharova T. I., Terentieev O. M. Decision support systems for modelling, forecasting and risk estimation: monography. Riga : LAP LAMBERT Academic Publishing. 2019. 176 p.
 99. Trofymchuk O. M., Bidiuk P. I., Prosyankina-Zharova T. I., Terentieev O. M. Decision support system for modeling and forecasting ecological processes. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях: колективна монографія* / ред. С. О. Довгий. Київ : Юстон, 2020. С. 23–44.
 100. Trofymchuk O. M., Bidiuk P. I., Prosyankina-Zharova T. I., Terentieev O. M. Forecasting nonstationary processes in demography, ecology, economy and finances. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях : колективна монографія* / ред. С. О. Довгий. Київ : Юстон, 2018. С. 113–123.

101. Box, G. E. P., Cox D. R. An Analysis of Transformations. *Journal of the Royal Statistics Society*. 1964. P. 211–252.
102. Ding F., Chenb T. Identification of Hammerstein nonlinear ARMAX systems. *Automatica*. 2005. №41. P. 1479–1489. URL: <http://www.paper.edu.cn/scholar/showpdf/NUT2EN2IMTD0ExeQh> (Last accessed: 18.09.2020).
103. Fung E. H., Wong Y. K., Ho H. F., Mignolet M. P. Modelling and prediction of machining errors using ARMAX and NARMAX structures. *Applied Mathematical Modelling*. 2003. Vol. 27, issue 8. P. 611–627. URL: <https://www.sciencedirect.com/science/article/pii/S0307904X03000714> (Last accessed: 19.03.2021).
104. Stoviček K. Forecasting with ARMA models. The case of Slovenian inflation. *Prikazi in analize*. 2007. Vol. XIV/1. P. 23–55.
105. Бокс Дж., Дженкинс Г. М. Анализ временных рядов. Москва : Мир, 1974. 288 с.
106. Griepentrog G. L., Ryan J. M., Smith, L. D. Linear Transformation of Polynomial Regression Models. *The American Statistician*. 1982. Vol. 36, issue 3a. P. 171–174.
107. White H. A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrics*. 1980. Vol. 48. P. 817–838.
108. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society*. 1996. Vol. 58, №1. P. 267–288.
109. De Jong P., Heller G. Z. Generalized Linear Models for Insurance Data. New York : Cambridge University Press, 2008. 197 p.
110. Hastie and Tibshirani *Generalized Additive Models*. New York: Chapman and Hall, 1990. 175 p.
111. McCullagh P., Nelder J. A. Generalized Linear Models. Second Edition. London : Chapman and Hall, 1989. 532 p.
112. Nelder J. A., Wedderburn R. W. M. Generalized Linear Models. *Journal of the Royal Statistical Society A*. 1972. Vol. 135. P. 370–384.

113. Abdi H. Partial least square regression (PLS regression). *Encyclopedia for research methods for the social sciences*. 2003. P. 792–795. URL: <https://www.utdallas.edu/~herve/Abdi-PLS-pretty.pdf> (Last accessed: 11.08.2020).
114. De Jong S. SIMPLS: An Alternative Approach to Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*. 1993. Vol. 18. P. 251–263.
115. Tobias R. D. An Introduction to Partial Least Squares Regression.” In Proceedings of the Twentieth Annual SAS Users Group International Conference, P. 1250–1257. Cary, NC: SAS Institute, Inc., 1995. URL: <http://support.sas.com/rnd/app/stat/papers/pls.pdf>. (Last accessed: 18.07.2020).
116. Bradley E., Hastie T., Tibshirani R., Johnstone Least Angle Regression. *Annals of Statistics*. 2004. Vol. 32 (2). P. 407–499.
117. Jim G., Jeff T., Chip W. Applied Analytics Using SAS Enterprise Miner: Course Notes. SAS Institute Inc. Cary, NC, USA, 2010. 698 p.
118. Amrhein J., Dickey D., Kiernan K., Marovich P. Statistics 2: ANOVA and Regression Course Notes. SAS Institute Inc. Cary, NC, USA. 2012. 712 p.
119. Freund R. J., Littell R. C. SAS System for Regression. 3 Edition. Cary, NC : SAS Institute, Inc., 2000. 256 p.
120. Iman R. The Analysis of Complete Blocks Using Methods Based on Ranks. *Proceedings of the SAS Users Group International Conference*. Cary, NC : SAS Institute, Inc., 1988. Vol. 13. P. 970–978.
121. Breiman L., Friedman J., Olshen R., Stone C. Classification and Regression Trees. Florida : CRC Press, 2000. 360 p.
122. Chow C. K., Liu C. N. Approximating discrete probability distributions with dependence trees. *IEEE Transactions on information theory*. 1968. Vol. IT-14, №3. P. 462–467.
123. Pashynska N., Snytyuk, V., Putrenko V., Musienko A. A decision tree in a classification of firehazard factors. *Восточно-Европейский журнал передових*

- технологий. 2016. №5(10). С. 32–37. URL: http://nbuv.gov.ua/UJRN/Vejpte_2016_5%2810%29_6 (дата обращения: 12.12.2020).
124. Quinlan J. R. Induction of decision trees. *Machine Learning*. Boston : Kluwer Academic Publishers. 1986. Vol. 1, № 1. P. 81–106.
 125. Quinlan J. R. C4.5: programs for machine learning. San Francisco : Morgan Kaufmann, 1993. 302 p.
 126. Galante R. Improving the Performance of Data Mining Models with Data Preparation Using SAS Enterprise Miner. *Proceeding SAS Global Forum*, April 26–29, 2015. Dallas, TX, USA, 2015. 21 p. URL: <https://support.sas.com/resources/papers/proceedings15/SAS1965-2015.pdf> (Last accessed: 11.08.2020).
 127. Adaptive fuzzy clustering of multivariates hort time series with unevenly distributed observations based on matrix neuro-fuzzy self-organizing network / Setlak G. et al. *Advances in Intelligent Systems and Computing*. 2018. Vol. 643. P. 308–315.
 128. Schubert S., Lee T. Time Series Data Mining with SAS Enterprise Miner. *The SAS Global Forum 2011 Conference*. 2011. 12 p. URL: <https://support.sas.com/resources/papers/proceedings11/160-2011.pdf>. (Last accessed: 12.09.2019).
 129. Gillies D. A. Lecture 11: Exact Inference / Course of lectures on «Intelligent data analysis and probabilistic inference». Imperial college of London : Department of computing, 2006. 4 p.
 130. Gilks W. R., Richardson S., Spiegelhalter D. J. Markov Chain Monte Carlo in Practice. New York : Chapman & Hall/CRC, 2000. 486 p.
 131. Bolstad W. M. Understanding computational Bayesian statistics. Hoboken (New Jersey): John Wiley & Sons, Ltd, 2010. 334 p.
 132. Top 10 algorithms in data mining / Wu X. et al. *Knowledge and Information Systems*. 2008. Vol. 14, №1. P. 1–37.

133. Tyshchenko O., Kopaliani D., Bodyanskiy Ye. The least squares support vector machines based on a neo-fuzzyneuron. *Computational Models for Business and Engineering Domains* / Eds. By G. Setlak, K. Markov. Rzeszow-Sofia : ITHEA. 2014. P. 44–51.
134. Bayes T. An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London*. 1763. Vol. 53. P. 370–418.
135. Cooper G., Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*. 1992. Vol. 9. P. 309–347.
136. Bernardo J. M., Smith A. F. M. Bayesian Theory. New York : John Wiley & Sons, Ltd, 2000. 586 p.
137. Hautaniemi S. K., Korpisaari P. T., Jukka P. P. Saarinen. Target identification with dynamic hybrid Bayesian networks. *Proceedings of the forth SPIE's international symposium on sensor fusion: architectures, algorithms and applications*. Orlando (Florida US). 2000. P. 55–66.
138. Heckerman D. Bayesian Networks for Data Mining. *Data Mining and Knowledge Discovery*. 1997. № 1. P. 79–119.
139. Heckerman D. E., Horvitz E. J., Nathwani B. N. Toward normative expert systems : Part I The Path Finder Project. *Methods of information in medicine*. 1992. Vol. 31, №2. P. 90–105.
140. Heckerman D., Geiger D., Chickering D. Learning Bayesian networks: the combination of knowledge and statistical data. *Technical report MSR-TR-94-09*. Microsoft Research. 1994. 53 p.
141. Бідюк П. І., Кузнецова Н. В. Основні етапи побудови і приклади застосування мереж Байєса. *Системні дослідження та інформаційні технології*. 2007. №4. С. 26–39.
142. Бідюк П. І., Литвиненко В. І., Кроптя А. В. Аналіз ефективності функціонування мережі Байєса. *Автоматика. Автоматизація. Електротехнічні комплекси та системи*. 2007. № 2. С. 6–15. : URL: http://nbuv.gov.ua/UJRN/aaeks_2007_2_3 (дата звернення: 31.01.2020).

143. Härdle W. K., Hautsch N., Mihoci A. Local adaptive multiplicative error models for high-frequency forecasts. *Journal of applied econometrics* 2015. Vol. 30. P. 529–550.
144. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. Москва : Финансы и статистика, 2003. 416 с.
145. Tsay R. S. Analysis of Financial Time Series. Hoboken : Wiley & Sons, Inc., 2010. 715 p.
146. Stanfill C., Waltz D. L. Toward memory-based reasoning. *Communications of the ACM*, 1986. Vol. 29, №12. P. 1213–1223.
147. Stepashko V. S., Efimenko S. N. Sequential Estimation of the Parameters of Regression Model. *Cybernetics and Systems Analysis*, 2005. Vol. 41. P. 631–634.
148. Leonard M., Sloan J., Lee ., Elsheimer B. An Introduction to Similarity Analysis Using SAS. *Proceedings of the SAS Global Forum 2007* . 2007. 22 p. URL: <https://support.sas.com/rnd/app/ets/papers/similarityanalysis.pdf>. (Last accessed: 10.03.2018).
149. Sankoff D., Kruskal J. Time warps, string edits, and macromolecules: The theory and practice of sequence comparison. Stanford, CA : CSLI Publications Stanford University Cordura, 1999. 408 p.
150. Gibbs B. P. Advanced Kalman Filtering, Least-squares and Modeling : a practical handbook. John Wiley & Sons, Inc., Singapore, 2011. 627 p. URL: <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470890042> (Last accessed: 10.01.2019).
151. Petersen I. R., Savkin A.V. Robust Kalman filtering for signals and systems with large uncertainties. Boston : Birkhouser USA, 1999. 214 p.
152. Балакришнан А. В. Теория фильтрации Калмана : пер. с англ. Москва : Мир, 1988. 168 с.
153. Axak N., Rosinskiy M. D. MapReduce Hadoop Models for Distributed Neural Network Processing of Big Data Using Cloud Services. *Conference on Computer Science and Information Technologies*. 2019. P. 387–400.

154. Bidyuk P., Prosyankina-Zharova T., Terentiev O., Medvedeva M. Adaptive modelling for forecasting economic and financial risks under uncertainty in terms of the economic crisis and social threats. *Технологічний аудит та резерви виробництва*. 2018. №4. С. 24–36. URL: [https://doi.org/ 10.15587/2312-8372.2018.135483](https://doi.org/10.15587/2312-8372.2018.135483).
155. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Москва : Финансы и статистика, 1986. 366 с.
156. Бард Й. Нелинейное оценивание параметров . Москва : Статистика, 1979. 349 с.
157. Бідюк П. І., Романенко В. Д., Тимошук О. Л. Аналіз часових рядів : навч. посіб. Київ : НТУУ КПІ, 2013. 600 с.
158. Freund R. J., Littell R. C. SAS System for Regression. 3 Edition. Cary, NC : SAS Institute, Inc., 2000. 256 p.
159. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды : пер. с англ. Москва : Наука, 1976. 736 с.
160. Кендэл М. Временные ряды : пер. с англ. Москва : Финансы и статистика, 1981. 199 с.
161. Меньяйленко О. С., Шевчук О. Б. Аналіз сучасних інформаційних експертних систем економічного напрямку. *Вісник Луганського національного університету імені Тараса Шевченка. Серія: Педагогічні науки*. Луганськ, 2012. № 21. С. 95–101. URL: http://nbuv.gov.ua/UJRN/vlup_2012_21_15 (дата звернення: 12.01.2021).
162. Stokes M. E., Davis C. S., Koch G. G. Categorical Data Analysis Using the SAS® System. 2 Edition. Cary, NC : SAS Institute Inc., 2000. 647 p.
163. Загальноекономічний і економіко-математичний словник Лопатнікова. URL: <https://lopatnikov.pro/?s=%D0%BD%D0%B5%D0%BB%D0%B8%D0%BD%D0%B5%D0%B9%D0%BD%D0%B0%D1%8F+%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0> (дата звернення: 17.02.2021).
164. Кембриджський словник. URL: <https://dictionary.cambridge.org/dictionary/english> (дата звернення: 11.01.2020).

165. Нелінійна система. URL: https://uk.wikipedia.org/wiki/Нелінійна_система#:~:text=Нелінійна%20система%20-%20динамічна%20система%2C%20в,займається%20дослідженням%20нелінійних%20динамічних%20систем.
166. Nonlinear system. URL: https://en.wikipedia.org/wiki/Nonlinear_system#:~:text=In%20mathematics%20and%20science%2C%20a,are%20inherently%20nonlinear%20in%20nature.
167. Новицький І. В., Ус С. А. Випадкові процеси : навч. посіб. Дніпропетровськ : Нац. гірничий ун-т, 2011. 125 с.
168. Cheng J., Bell D. A., Liu W. Learning belief networks from data: an information theory based approach. *Proceedings of the sixth international conference on information and knowledge management (CIKM 1997)*. Las Vegas (Nevada), 1997. P. 325–331.
169. Mirollo R. E., Strogatz S. H. Amplitude death in an array of limit-cycle oscillators. URL: <https://link.springer.com/article/10.1007/BF01013676?LI=true> (Last accessed: 01.04.2020)
170. Енциклопедія математики. URL: <https://encyclopediaofmath.org/index.php?title=Chaos> (дата звернення: 15.03.2021)
171. Boussinesq J. Theorie de l'intumescence Liquid, Appleteonde Solitaire au de Translation, se Propageant dans un Canal Rectangulaire. *Les Comptes Rendus de l'Académie des Sciences*, 1871. Vol. 72. P. 755–759.
172. Jenkins A. Self-oscillation. *Physics Reports*. 2013. 525 (2). P. 167–222. DOI: [10.1016/j.physrep.2012.10.007](https://doi.org/10.1016/j.physrep.2012.10.007).
173. Box G. E. Non-normality and Tests on Variance. *Biometrika*. 1953. №40. P. 318–335.
174. Dickey D. A., Fuller W. A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*. 1979. №74 (366). P. 427 – 431.
175. Бідюк П. І., Терентьев О. М., Просянкін-Жарова Т. І. Прикладна статистика : навч. посіб. Вінниця : Едельвейс і К, 2013. 288 с.

176. Siddiqi N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring. Wiley and SAS Business, USA, 2015. 208 p.
177. Thode H. C. Jr. Testing for Normality. New York : Marcel Dekker, Inc., 2002. 495 p.
178. Демиденко Е. З. Линейная и нелинейная регрессия. Москва : Финансы и статистика, 1981. 302 с.
179. Гожий О. П. Підхід до оцінювання невизначеностей в задачах прогнозування. *Електротехнічні та комп'ютерні системи*. 2015. № 19(95). С. 243–247.
180. Згуровський М. З., Панкратова Н. Д. Основи системного аналізу. Київ : Видавнича група BHV, 2007. 544 с.
181. Payzan-LeNestour E., Bossaerts P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian Learning in Unstable Settings. *PLoS Computational Biology*. 2011. Vol. 7, issue 1. P. 1–14.
182. Chatfield C. Model Uncertainty, Data Mining and Statistical Inference. *Journal of the Royal Statistical Society*. 1995. № 158. P. 419–466.
183. Rykov A. C. System analysis methods: multi-criteria and fuzzy optimization, modeling and expert estimates. Moscow : Economics, 1999. 1587 p.
184. Das B. Representing uncertainties using Bayesian Networks / Technical report DSTO-TR-0918, Defence Science and Technology Organization, Electronics and Surveillance Research Laboratory. Australian Department of Defense, 1999. 66 p.
185. Penman S. H. Financial forecasting, risk, and valuation: accounting for the future. *Abacus*. 2010. Vol. 46, issue 2. P. 211–228
186. McNeil A. J., Frey R., Embrechts P. Quantitative Risk Management. – Princeton. New Jersey : Princeton University Press, 2005. 538 p.
187. Козьменко О. В., Кузьменко О. В. Економіко-математичні методи та моделі (економетрика) : навч. посіб. Суми : Університетська книга, 2014. 406 с.

188. Ларичев О. И., Мошкович Е. М. Качественные методы принятия решений. Вербальный анализ решений. Москва : Наука. Физматлит, 1996. 208 с.
189. Shahar Y. Computing with Influence Diagrams and the Path Finder Project : Course of lectures on “Judgment and Decision Making in Information Systems”/ Ben-Gurion University of the Negev, department of Information System Engineering. 2008. Lecture 6. 26 p.
190. Trofymchuk O., Bidyuk P. Probabilistic and statistical uncertainty Decision Support Systems. *Visnyk of Lviv Polytechnic National University*. 2015. No. 826. P. 237–248.
191. Трофимчук О., Бідюк П., Кожухівська О., Кожухівський А. Ймовірнісно-статистичні невизначеності в системах підтримки прийняття рішень. *Вісник Національного університету «Львівська політехніка»*. 2015. № 826. С. 237–248.
192. Galante R. Improving the Performance of Data Mining Models with Data Preparation Using SAS Enterprise Miner. *Proceeding SAS Global Forum*, April 26–29, 2015. Dallas, TX, USA, 2015. 21 p. URL: <https://support.sas.com/resources/papers/proceedings15/SAS1965-2015.pdf> (Last accessed: 11.08.2020).
193. Matignon R. Data-mining using SAS Enterprise Miner. 2007. 584 p. URL: <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470149019.html>
194. Enders C.K. Applied Missing Data Analysis. NY : The Guilford Press, 2010. 401 p.
195. Multiple imputation in SAS / Institute for digital research and education. 2016 URL: http://stats.idre.ucla.edu/sas/seminars/multiple-imputation-in-sas/mi_new_1 (Last accessed: 11.08.2020).
196. Zhang Zhongheng. Missing data imputation: focusing on single imputation. *Annals of Translational Medicine*. 2016. №4(1), 9. P. 11–18. DOI: 10.3978/j.issn.2305-5839.2015.12.38.
197. Olson D.L., Delen D. Advanced Data Mining Techniques. 2008. 181p.

198. Айвазян С. А. Теория вероятности и прикладная статистика. Москва : Юнити, 2001. 656 с.
199. Андерсен Т. Введение в многомерный статистический анализ : пер. с англ. Москва : Физматгиз, 1963. 500 с.
200. Андерсон Т. Статистический анализ временных рядов : пер. с англ. Москва : Мир, 1976. 755 с.
201. Гмурман В. Е. Теория вероятностей и прикладная статистика. Москва : Высшая школа, 2002. 479 с.
202. Xekalaki E., Degiannakis S. ARCH models for financial applications. Padstow, Cornwall, Greate Britain : TJ International, 2010. 558 p.
203. Seber G. A. F., Wild C. J. Nonlinear Regression. New York : John Wiley & Sons, Inc, 1989. 800 p.
204. Neter J., Kutner M. H., Wasserman W., Nachtsheim C. J. Applied Linear Statistical Models. Fourth Edition. New York : WCB McGraw Hill, 1996. 1424 p.
205. Hurvich C. M., Simonoff J S., Tsai C. L. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society*. 1998. P. 271–293.
206. Anderson W. N., Kleindorfer G. B., Kleindorfer P. R., Woodrooffe M. B. Consistent estimates of the parameters of a linear system. *The Annals of Mathematical Statistics*. 1969. Vol. 40, №6. P. 2064–2075.
207. Bowerman B. L., O'Connell R. T., Dickey D. A. Linear Statistical Models, an Applied Approach. Boston, MA: Duxbury Press, 1986. 1024 p.
208. Chartrand, R., Yin, W. Iteratively reweighted algorithms for compressive sensing. *IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2008. P. 3869-3872.
209. Iman R. Some Aspects of the Rank Transform in Analysis of Variance Problems. *Proceedings of the SAS Users Group International Conference*. Cary, NC: SAS Institute, Inc., 1982. Vol. 7. P. 676–680
210. Peixoto J. L. Hierarchical Variable Selection in Polynomial Regression Models. *The American Statistician*. 1987. Vol. 41, No. 4. P. 311–313.

211. Miller D. M. Reducing Transformation Bias in Curve. Fitting. *The American Statistician*. 1984. Vol. 38, 2. P. 124–126.
212. Olejnik S. F., Algina J. Type I Error Rates and Power Estimates of Selected Parametric and Non-parametric Tests of Scale. *Journal of Educational Statistics*. 1987. Vol. 12. P. 45–61.
213. Searle S. R. Linear Models for Unbalanced Data. New York : John Wiley & Sons, Inc, 1987. 551 p.
214. SAS Institute Inc. Create a gradient boosting model of the data. Getting started with SAS® Enterprise Miner™ 7.1. URL: [//support.sas.com/documentation/cdl/en/emgsj/64144/HTML/default/viewer.htm#p03iy98sk0c9bvn1r6x7ppx8uj08.htm](http://support.sas.com/documentation/cdl/en/emgsj/64144/HTML/default/viewer.htm#p03iy98sk0c9bvn1r6x7ppx8uj08.htm) (Last accessed: 11.08.2020).
215. Huber P. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*. 1964. Vol. 35(1). P. 73-101
216. Goodall C. M-estimators of location: An outline of the theory. *Understanding Robust and Exploratory Data Analysis*. New York, 1983. P. 339-403.
217. Andrews' Wave. StatWiki internet resource. URL: http://141.20.100.229/mediawiki/statwiki/index.php/Andrews%E2%80%99_Wave. (Last accessed: 21.02.2020).
218. Martin A. T. The Data Augmentation Algorithm. *Tools for Statistical Inference*. New York, Inc : Springer-Verlag, 2011. P. 90–136. URL: https://link.springer.com/chapter/10.1007%2F978-1-4612-4024-2_5 (Last accessed: 13.05.2020).
219. Бідюк П. І. Терентьев О. М., Просьянкіна-Жарова Т. І., Савастьянов В. В. Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни. *International conference on System Analysis and Information Technology (SAIT 2017) : Proceedings of the 19-th. Conference (Kyiv, May 22–25 2017)*. Kyiv : National Technical University of Ukraine “KPI”, 2017. С. 253–255.

220. Бідюк П. І., Терентьєв О. М., Просянкін-Жарова Т. І. Методи заповнення пропусків даних у задачах прогнозного моделювання соціально-економічних процесів. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2016)* : матеріали міжнар. конф. (м. Залізний порт, 22–26 травня 2017 р.). Херсон : Вишемирський В. С., 2017. С. 185–187.
221. SAS Institute Inc. SAS/STAT User's Guide. Version 9.1. Cary, NC : SAS Institute, Inc., 2004. 5136 p.
222. Державна служба статистики України : офіц. сайт. URL: <http://www.ukrstat.gov.ua/> (дата звернення: 12.01.2020).
223. Resource potential of agrarian sphere: problems and tasks of effective use / Sobkevich. V. et al. Kyiv : NISS, 2013. ...p.
224. Kleinbaum D. G., Kupper L. L., Muller K. E. Applied Regression Analysis and Other Multivariable Methods. Boston, MA : PWS-Kent Publishing Company, 1988. 718 c.
225. Kotz S., Johnson N. L., Read C. B. Encyclopedia of Statistical Sciences. New York : John Wiley & Sons, 1988. 336 p.
226. Ratkowsky D. Handbook of Nonlinear Regression Models. New York; Basel : Marcel Dekker, 1990. 241 p.
227. Peixoto J. L. A Property of Well-Formulated Polynomial regression Models. *The American Statistician*. 1990, Vol. 44, No. 1. P. 26–30.
228. Jordan M. I., Ghahramani Z. G., Jaakkola T. S., Saul L. K. An introduction to variational methods for graphical models. *Machine Learning*. Cambridge : MIT Press, 1999. Vol. 37, №2. P. 183–233.
229. Anscombe F. Graphs in Statistical Analysis. *The American Statistician*. 1973. №27. P. 17–21.
230. Miller R. G., Jr. Beyond ANOVA: Basics of Applied Statistics. Boca Raton, FL : Chapman & Hall/CRC. 1997. 336 p.

231. O'Brien R. G. A General ANOVA Method for Robust Tests of Additive Models for Variances. *Journal of the American Statistical Association*. 1979. Vol. 74. P. 877–880.
232. Belsey D. A., Kuh E., Welsch R. E. Regression Diagnostics: Identifying Influential data and sources of collinearity. New York : John Wiley & Sons, 1980. XV, 292 p.
233. Shergin V. L., Chala L. E., Udovenko S. G. Fractal dimension of infinitely growing discrete sets. *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*. 2018. P. 259–263.
234. Milliken G. A., Johnson D. E. Analysis of Messy Data, Volume 1: Designed Experiments. New York : Chapman & Hall, 1992. 690 p.
235. Agresti A. An Introduction to Categorical Data Analysis. New York: John Wiley & Sons, 1996. 718 p.
236. Chakraborty G., Pagolu M. Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS. Cary, NC: SAS Institute Inc. 2013. 564 p.
237. Art D., Gnanadesikan R., Kettenring, R. Data-based Metrics for Cluster Analysis. *Utilitas Mathematica*. 1983. Vol. 21. P. 75–99.
238. Gorodetsky V. I., Drozdgin V. V., Jusupov R. M. Application of attributed grammar and algorithmic sensitivity model for knowledge representation and estimation. *Artificial Intelligence and Information, Control Systems of ROBOTS*. North Holland : Elsevier Science, 1984. P. 232–237.
239. Терентьев О. М., Просьянкіна-Жарова Т. І., Бідюк П. І., Связинська Н. О., Кириченко В. Е. Text mining analysis of agriculture internet sources using SAS software. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI 2016)* : матеріали міжнар. конф. (м. Залізний порт, 24–28 трав. 2016 р.). Херсон : ХНТУ, 2016. С. 244–246.
240. Терентьев О. М., Просьянкіна-Жарова Т. І., Савастьянов В. В. Використання засобів текстової аналітики як інструменту оптимізації підтримки

- прийняття рішень у задачах розробки планів соціально-економічного розвитку України. *Реєстрація, зберігання і обробка даних*. 2016. Т. 18. № 3. С. 75–86.
241. Шолохов О. В., Терентьев О. М. Просьянкіна-Жарова Т. І. Підвищення ефективності соціальних комунікацій на основі аналізу інтернет-джерел засобами textmining. *Прикладні системи та технології в інформаційному суспільстві* : матеріали міжнар. наук.-практ. конф. (Київ, 30 верес. 2020 р.). Київ : Київський нац. ун-т імені Тараса Шевченка, 2020. С. 23–44.
242. Бідюк П. І., Коршевнік Л. О., Терентьев О. М. Підтримка розв’язання слабкоструктурованих задач в органах державної влади. *Системный анализ и информационные технологии (САИТ-2012)* : матеріали 14 міжнар. наук.-техн. конф. (Київ, 24 квіт. 2012 р.). Київ : ННК “ІПСА” НТУУ “КПІ”. 2012. С. 169–170.
243. Wilbur W. J., Lipman D. J. Rapid similarity searches of nucleic acid and protein data banks. PubMed. URL: <https://pubmed.ncbi.nlm.nih.gov/6572363/> (Last accessed: 18.07.2020).
244. Клейнберг Дж., Тардос Е. Алгоритмы: разработка и применение. Классика Computers Science. СПб.: Питер, 2016. 800 с.:
245. Andreassen S., Woldbye M., Falck B., Andersen S. K. MUNIN – a causal probabilistic network for interpretation of electromyographic findings. *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*. Los Altos, CA : Morgan Kaufmann Publishers, 1987. P. 366–373.
246. Dempster A. P., Laird N. M., Rubin D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1977. Vol. 39, No. 1, P.1-38.
247. Diez F. J., Mira J., Iturralde E., Zubillaga S. DIAVAL, a Bayesian expert system for echocardiography. *Artificial Intelligence in Medicine*. 1997. Vol. 10. P. 59–73.

248. Falzon L. Using Bayesian network analysis to support centre of gravity analysis in military planning. *European Journal of Operational Research*. 2006. Vol. 170, № 2. P. 629–643.
249. Friedman N., Linia M., Nachman I. Using Bayesian networks to analyze gene expression data. *Journal of Computational Biology*. 2000. Vol. 7. P. 601–620.
250. Friedman N., Ninio M., Pupko T. A structural EM algorithm for phylogenetic inference. *Journal of Computational Biology*. 2002. Vol. 9. P. 331–353.
251. Fung R., Favero B. Applying Bayesian Networks to information retrieval. *Communications of the ACM (CACM)*. 1995. Vol. 38, № 3. P. 24–30.
252. Galan S. F., Aguado F., Diez F. J., Mira J. NasoNet, Joining Bayesian Networks and Time to Model Nasopharyngeal Cancer Spread. *AIME*, Heidelberg : Springer-Verlag, 2001. P. 207–216.
253. Gelman A., Carlin J. B., Stern H. S., Rubin D. B. Bayesian Data Analysis. New York : Chapman and Hall, CRC Press, 2000. 670 p.
254. Guo H. Probabilistic inference in Bayes networks / Report on research activities, Kansas State University, department of computing and information sciences, November 2000. 27 p.
255. Guo H. Real time Bayesian networks inference / Report on research activities. Kansas State University, department of computing and information sciences, March November 2001. 42 p.
256. Jensen F.V., Nielsen Th.D. Bayesian Networks and Decision Graphs. New York : Springer, 2007. 457 p.
257. Jilani T., Najamuddin M. Adaptive Bayesian modeling for time series forecasting. *Journal of Applied Environmental and Biological Sciences*. 2014. Vol. 4(7S). P. 99–106.
258. Jouffe L., Munteanu P. New search strategies for learning Bayesian networks. *Proceedings of tenth international symposium on applied stochastic models and data analysis (ASMDA 2001)*. Compiègne (France). 2001. Vol. 2. P. 591–596.

259. Kim J. H., Pearl J. CONVINCe: A conversational inference consolidation engine. *IEEE Transactions on systems, Man and Cybernetics*. 1983. Vol. 17(2). P. 120–132.
260. Kozlov A.V., Singh J. P. A parallel Lauritzen-Spiegelhalter algorithm for probabilistic inference. *Proceedings of the Supercomputing'94 conference*. Washington, DC, November, 1994. P. 320–329.
261. Lauritzen S. L., Spiegelhalter D. J. Local computations with probabilities on graphical structures and their application to expert systems. *Journal Royal Statistics Society, series B – statistical methodology*. 1988. Vol. 50, №2. P. 157–194.
262. Lo B., Thiemjarus S., Yang G. Adaptive Bayesian networks for video processing. *IEEE International conference on image processing (ICIP)*, Barcelona (Spain). 2003. 14–17 September. №1. P. 889–892.
263. Mani S., McDermott S., Valtorta M. MENTOR: a Bayesian Model for prediction of mental retardation in newborns. *Research in Developmental Disabilities*. 1997. Vol. 18, №5. P. 303–318.
264. Neil M., Fenton N. E., Tailor M. Using Bayesian networks to model expected and unexpected operational losses. *Risk Analysis*. 2005. Vol. 25, No. 4. P. 34–57.
265. Onisko A., Druzdzel M. J., Wasyluk H. A probabilistic causal model for diagnosis of liver disorders. *Proceedings of the Seventh Symposium on Intelligent Information Systems (IIS-98)*. Malbork, Poland, 1998. P. 379–387.
266. Pearl J. Causality: models, reasoning, and inference. Cambridge University Press, 2000. 400 p.
267. Pourret O., Naim P., Marcot B. Bayesian Networks: A Practical Guide to Applications. Chichester, UK : Wiley, 2008. 448 p.
268. Probabilistic diagnosis using a reformulation of the INTERNIST- 1/QMR knowledge base I. The probabilistic model and inference algorithms / Shwe M. et. al. *Methods of Information in Medicine*. 1991. Vol. 30(4). P. 241–255.

269. Rossi P., Allenby G. M. Bayesian Statistics and Marketing. Hoboken. New Jersey : Wiley & Sons, Ltd., 2005. 350 p.
270. Santos E. J., Shimony S. E., Solomon E., Williams E. On a distributed anytime architecture for probabilistic reasoning : technical report TR-95-02, department of electrical and computer engineering / Air Force Institute of Technology. Ohio, USA : Wright-Patterson AFB, 1995. 18 p.
271. Shenoy P., Shenoy P. Bayesian Network models of portfolio risk and return. *Proceedings of the sixth annual international conference Computational Finance 1999* / New York University's Stern School of Business. Cambridge : MIT Press, 2000. P. 87–106.
272. Sierra B, Larranaga P. Predicting survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. *Artificial Intelligence in Medicine*. 1998. Vol. 14. P. 215–230.
273. Spirtes P., Glymour C., Scheines R. Causation, prediction and search. *Adaptive computation and machine learning*. MIT press, 2001. 565 p.
274. Learning Bayesian networks from data: an information-theory based approach / Cheng J. et al. *The artificial intelligence journal (AIJ)*. 2002. Vol. 137. P. 43–90.
- Littell R. C., Stroup W. W., Freund R. J. SAS® System for Linear Models. Cary, NC : SAS Institute Inc, 2002. 487 p.
275. Golmard J. L., Mallet A. Learning probabilities in causal trees from incomplete databases. *Revue d'Intelligence Artificielle*. 1991. Vol. 5. P.93–106.
276. Rebane G., Pearl J. The recovery of causal poly-trees from statistical data. // *International journal of approximate reasoning*. 1988. Vol. 2, № 3. P. 175–182.
277. Гасанова Л. Т., Терентьев О. М., Мельник І. В. Застосування мережі Байеса в медицині. *Вісник Університету "Україна"*. Київ : Університет "Україна", 2008. № 6. С. 93–96.

278. Бидюк П. И., Терентьев А. Н. Метод вероятностного вывода в байесовских сетях по обучающим данным. *Кибернетика и системный анализ*. 2007. № 3. С. 93–99.
279. Згуровський М. З., Бідюк П. І., Терент'єв О. М. Системна методика побудови байєсових мереж. *Наукові вісті НТУУ "КПІ"*. 2007. №4. С. 47–61.
280. Згуровський М. З., Бідюк П. І., Терент'єв О. М., Просянкіна-Жарова Т. І. Байєсівські мережі в системах підтримки прийняття рішень : навч. посіб. Київ : Едельвейс, 2015. 300 с.
281. Бідюк П. І., Терент'єв О. М., Коновалюк М. М. Байєсівські мережі в технологіях інтелектуального аналізу даних. *Наукові праці [Чорноморського державного університету імені Петра Могили комплексу "Києво-Могилянська академія"]*. Серія:Комп'ютерні технології. 2010. Т. 134. Вип. 121. С. 6–16.
282. Терент'єв А. Н., Бидюк П. И., Коршевніук Л. А. Алгоритм вероятностного вывода в байесовских сетях. *Системні дослідження та інформаційні технології*. 2009. № 2. С. 107–111
283. Bidyuk P. I., Davidenko V. I., Trofimenko D. V., Terentyev A. N. Comparative analysis of estimation methods of vertices correlation while Bayesian networks construction. *Journal of Automation and Information Sciences*, 2010. №42(11). P. 36–45. (Scopus)
284. Prosyankina-Zharova T. I., Terentiev O. M., Bidyuk P. I., Gasanov A. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and forecasting. *On control and optimization with industrial applications (COIA-2015) : Proceedings of the fifth international conference (Baku, 27–29 August 2015)*, Baku : Institute of applied mathematics BSU, 2015. P. 365–368.
285. Prosyankina-Zharova T., Terentiev O., Bidyuk P., Makukha M. Features of SAS Enterprise Guide for probabilistic modeling system, macroeconomic analysis and forecasting. *Journal of Mathematics and System Science*. 2016. №6. P. 112–122. (Index of Copurnicus)

286. Бідюк П. І., Коршевніюк Л. О., Терентьєв О. М., Просянкін-Жарова Т. І. Аналіз інвестиційних і соціально-економічних процесів методами моделювання обмежених множин багатовимірних даних. *Наукові вісті КІІІ*. 2012. № 2. С. 87–93.
287. Собкевич О. В., Русан В. М., Юрченко А. Д., Ковальова О. В. Ресурсний потенціал аграрної сфери: проблеми та завдання ефективного використання. Київ : НІСД, 2013. 76 с.
288. Обидєннова Т. С. Використання моделей та методів для формування та прийняття ефективних управлінських рішень. *Вісник Харківського національного технічного університету сільського господарства імені Петра Василенка*. Харків, 2016. Вип. 172. С. 147–153. URL: http://nbuv.gov.ua/UJRN/Vkhdtusg_2016_172_18 (дата звернення: 21.12.2020)
289. Бідюк П. І., Кириченко В. Е., Свєзінська Н. О. Комп'ютерна програма “IMLBayesNet” : авторське свідоцтво № 64117 України ; заявл. 17.12.2015 № 64562 ; опубл. 16.02.2016.
290. Трофименко Д. В., Давиденко В. І. Комп'ютерна програма “Bayesian Network Master BNetMaster” : авторське свідоцтво № 34443 України ; заявл. 09.06.2010; № 34657 ; опубл. 09.08.2010.
291. Малахов Е. В., Востров Г. Н. Интеллектуальные системы информационной поддержки процедур принятия решений. *Материалы междунар. науч. конф. «Искусственный интеллект – 2000»*. Таганрог : ТРТУ, 2000. С. 40–41.
292. Littell R. C., Milliken G. A., Stroup W. W., Wolfinger R. D. SAS® System for Mixed Models, Cary, NC : SAS Institute Inc., 1996. 828 p.
293. Довгий, С. О., Бідюк П. І., Трофимчук О. М. Системи підтримки прийняття рішень на основі статистично-ймовірнісних методів. Київ : Логос, 2014. 418 с. : рис., табл.
294. Modeling and forecasting financial and economic processes with decision support system / Bidyuk P. I. et al. *Наукові вісті КІІІ*. 2019. № 5–6. С. 7–17. URL: <https://doi.org/10.20535/kpi-sn.2019.5-6.176835>

295. Свиридчук Я. В., Дегтярьов А. О. Терентьев О. М. Застосування нечіткого множинного методу для оцінювання ризику банкрутства корпорації. *Вісник Університету «Україна». Серія: Інформатика, обчислювальна техніка та кібернетика*. 2010. № 8. С. 105–108.
296. Korbicz J., Bidyuk P., Kuznietsova N., Terentiev O., Prosiiankina-Zharova T. Multivariate distribution model for financial risks management. *The 9th International Conference "Information Control Systems & Technologies" (ICST2020)* (Odessa, Sept. 24–26, 2020). CEUR Workshop Proceedings, 2020. P. 416–429. (Scopus)
297. Terentyev A. N., Bidyuk P. I., Mironova A. V., Medin N. Y. Comparison of data mining methods while credit rating of natural persons. *Journal of Automation and Information Sciences*, 2009. № 41(10). P. 71–80. (Scopus)
298. Терентьев О. М., Бідюк П. І., Кузнєцова Н. В. Система підтримки прийняття рішень для аналізу фінансових даних. *Наукові вісті КНУ*. 2011. №1. С. 48–61.
299. Космічний простір на службі аграрія URL: <http://www.agroprofi.com.ua/statti/1115-kosmichnij-prostir-na-sluzhbi-agrarija.html>. (дата звернення: 23.01.2018).
300. Моніторинг навколишнього середовища з використанням космічних знімків супутника NOAA : монографія / Довгий С. О. та ін. ; Ін-т телекомунікацій і глобал. інформ. простору НАН України, Нац. аерокосм. ун-т ім. М. Є. Жуковського "ХАІ". Київ : Пономаренко Є. В., 2013. 316 с. : іл.
301. Регрессионные модели оценки урожайности сельскохозяйственных культур по данным MODIS / Н. Н. Куссуль и др. *Современные проблемы дистанционного зондирования Земли из космоса*. 2012. Т. 9, № 1. С. 95–107.
302. Малахов Е. В., Лисенко М. І., Полін Є. Л. Організація систем інформаційного обігу між картографічними системами та банками даних. *Матеріали 35-ї наукової конференції молодих дослідників ОПУ* –

- магістрантів «Сучасні інформаційні технології та телекомунікаційні мережі». Одеса : ОДПУ, 2000. С. 33.
303. Трофимчук О. М., Бідюк П. І., Просянкін-Жарова Т. І., Терент'єв О. М. Адаптивне моделювання і прогнозування у системах підтримки прийняття рішень сільськогосподарського призначення. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : колективна монографія / ред. С. О. Довгий. Київ : Юстон, 2017. С. 21–28.
304. Бідюк П. І., Терент'єв О. М., Просянкін-Жарова Т. І., Ефендієв В. В. Прогнозне моделювання нелінійних нестационарних процесів у рослинництві з використанням інструментів SAS Enterprise Miner. *Наукові вісті НТУУ КПП*. 2017. №1. С. 24–36. URL: <https://doi.org/10.20535/1810-0546.2017.1.87423>.
305. Терент'єв О. М., Просянкін-Жарова Т. І., Саваст'янов В. В. Застосування когнітивного та ймовірнісного моделювання у задачах формування сценаріїв розвитку соціально-економічних систем. *Наукові вісті НТУУ КПП*. 2016. №5. С. 37–47. URL: <https://doi.org/10.20535/1810-0546.2016.5.79876>.
306. Research semistructured problems of socio-economic systems: cognitive approach / Gorelova G.V. et al. Rostov na Donu : Publishing House of Rostov. University Press, 2006. 332 p.
307. Terentiev O. M., Prosiankina-Zharova T. I., Lahno V. A., Usatiuk Y. V. The features of the predictive computing modeling power system load in terms of reforming energy market. *Journal of Theoretical and Applied Information Technology*. 2020. Vol. 98. No. 2. P. 163–182. (Scopus Q3)
308. Терент'єв О. М., Св'язінська Н. О., Просянкін-Жарова Т. І. Візуалізація типології регіонів України із використанням система SAS Enterprise Guide 7.1. *Геоинформационные системы и компьютерны етехнологии эколого-экономического мониторинга* : матеріали міжнар. наук.-техн. конф. (Дніпропетровськ, 13–15 квітня 2016 р.). Днепропетровск : ГВУЗ «НГУ» МОН Украины, 2016. С. 21–25.

309. Akaike H. Information theory and an extension of the maximum likelihood principle. *Proceedings of the 2nd International Symposium on Information Theory* / eds: Petrov, Csaki. 1973. P. 267–281.
310. Mallows C. L. Some Comments on Cp. *Technometrics*. 1973. Vol. 15. P. 661–675 p.
311. Power D. J. What is a DSS? *The On-Line Executive Journal for Data-Intensive Decision Support*. 1997. Vol. 1., N3. P. 46–54.
312. Системи баз даних та знань. Кн. 2 : Системи управління базами даних та знань : навч. посіб. Львів : Магнолія-2006, 2008. 584 с.
313. Kondratenko Y.P., Kondratenko G. V., Sidenko I. V. (2018) Knowledge-Based Decision Support System with Reconfiguration of Fuzzy Rule Base for Model-Oriented Academic-Industry Interaction. *Applied Mathematics and Computational Intelligence*. FIM 2015 / eds: Gil-Lafuente A., Merigó J., Dass B., Verma R. *Advances in Intelligent Systems and Computing*. Cham : Springer, 2018. Vol. 730. P. 101–112.
314. Моделі оптимізації багаторівневого зберігання даних / Теленик С. Ф. та ін. *Вісник НТУУ «КПІ»*. Серія: Інформатика, управління та обчислювальна техніка : зб. наук. праць. 2015. Вип. 63. С. 48–53.
315. Берко А. Ю., Верес О. М., Пасічник В. В. Системи баз даних та знань. Кн. 1. Львів : Магнолія-2006, 2008. 440 с.
316. Бідюк П. І., Демківський Є. О. Проектування систем підтримки прийняття рішень : навч. посіб. Київ : КНУТД, 2005. 156 с.
317. Малахов Е. В. Вопросы организации иерархических информационных хранилищ. *Перспективы*. 1997. № 1. С. 122–123.
318. SAS Energy Forecasting 3.2: User's Guide, Second Edition. Cary, NC : SAS Institute Inc., 2017. 302 p. URL: <https://support.sas.com/documentation/onlinedoc/ef/3.2/efug.pdf> (*Last accessed:*
319. SAS Energy Forecasting: Fact Sheet. 2018. 4 p. URL: sas.com/energy-forecasting

320. Домрачев В. Н., Костецкий Р. И. Терентьев А. Н. SAS BASE: Основы программирования. Киев : Эдельвейс, 2014. 304 с.

ДОДАТКИ

ДОДАТОК А

Акт
про впровадження результатів дисертаційного дослідження
Присянкіної-Жарової Тетяни Іванівни
на тему «Інформаційні технології для систем підтримки прийняття
рішень в управлінні регіональним розвитком»

За результатами розгляду матеріалів дисертаційного дослідження встановлено, що представлені в роботі інформаційна технологія підтримки прийняття рішень, основу яких становлять методи інтелектуального аналізу даних, методика сценарного аналізу, математичні, економетричні, когнітивні моделі та ансамблі моделей є актуальними та мають практичну цінність.

Адже по результатам дослідження науково-прикладної теми можливо зробити висновки про доцільність та перспективність використання запропонованих рішень в різних сферах державного управління, у тому числі в наступному:

- Цивільний захист населення (особливо важливо під час воєнного стану) - сценарне планування (у тому числі динамічне планування використовуючи наявну оперативну інформацію з місць):
 - При надзвичайних ситуаціях зв'язаних з підтопленням, у тому числі внаслідок порушення цілісності гребель.
 - При надзвичайних ситуаціях техногенного характеру (наприклад аварії на хімічному виробництві)
- Стратегії регіональних (обласних) рівнів:
 - Підготовка інформації для більш якісного SWOT-аналізу та PEST-аналізу, враховуючи тенденції у тому числі соціальні (text-mining аналіз)
 - Розробка сценаріїв (з метою вибору базового сценарію):
 - «Як є» - зазвичай такі сценарії розробляються на базі визначених тенденцій і є наглядною інформацією що буде якщо нічого не робити.
 - Оптимістичний, реалістичний та песимістичний сценарії
 - Аналіз вірогідності досягнення стратегії на базі проміжної звітної інформації

В цілому використання досліджених у науково-прикладній темі засобів сучасних інформаційних технологій та прогнозного моделювання підвищать ефективність підтримки прийняття рішень в управлінні розвитком складних соціально-економічних систем – регіонів. Це дозволить удосконалити процес розроблення регіональних стратегій, цільових та інвестиційних програм, програм інформатизації, тощо, підвищити інформативність моніторингу їх виконання та аналізу результатів.

Результати апробації інформаційної технології як в цілому, так і окремих моделей та методів моделювання, аналізу, оброблення структурованих та неструктурованих даних, даних призначені для прогнозування показників соціально-економічного розвитку регіону, дозволяють зробити висновки щодо

Черкаська обласна державна адміністрація
№ 01/01-21/10417/01/01-21/23311 від 04.09.2025



перспективності їх використання на рівні регіонів та припущення про можливе застосування на загальнодержавному рівні в подальшому.

Як перший крок для практичного впровадження результатів дослідження в умовах реального управління, можливо відпрацювання спільного рішень з вже існуючими інформаційними системами в країні, наприклад системами Ситуаційної обізнаності великих міст.

Даний документ не є підставою для проведення розрахунків та передачі майнових чи авторських прав, виникнення будь-яких, в тому числі фінансових зобов'язань.

Заступник голови ЧОДА з питань цифрового
розвитку, цифрових трансформацій і цифровізації

Сергій СВЕРЧКОВ

Начальник Відділу з питань цифрового розвитку,
цифрових трансформацій та цифровізації

Дмитро ХУЛАП

АКТ

впровадження результатів дисертаційного дослідження
Тетяни Іванівни Просянкіної-Жарової на тему:
«Інформаційні технології для систем підтримки прийняття рішень в
управлінні регіональним розвитком»

Агропромислове виробництво – основа економіки більшості регіонів країни, особлива сфера, яка не лише є джерелом наповнення місцевого бюджету, а й осередком, навколо якого гуртується громада. Тому питання визначення шляхів забезпечення його ефективності та сталого розвитку, важливе як на рівні підприємства, та як і для територіальної громади та регіону. Одним із варіантів розв'язання даної проблеми є використання спеціалізованих систем підтримки прийняття рішень, у яких реалізовано підсистему прогнозного моделювання урожайності сільськогосподарських культур з можливістю інтелектуального та просторового аналізу, обробки зображень, картографування.

Тому, результати даного дисертаційного дослідження мають практичне значення й були застосовані у роботі підприємства. Апробацію представленої в роботі підсистеми системи прогнозування урожайності сільськогосподарських культур, виконано на прикладі галузі рослинництва підприємства для ділянок, що мають різний екологічний стан. Вхідними наборами даних використані картографічні дані, статистика урожайності сільськогосподарських культур, значення індексу NDVI та показників стану навколишнього середовища.

Результати апробації підсистеми в цілому, окремих моделей та методів аналізу часових рядів, моделювання, обробки геопросторових даних, призначених для прогнозування урожайності сільськогосподарських культур, виявлення чинників, що найбільше впливають на підвищення ефективності рослинництва, дозволяють зробити висновки щодо доцільності використання результатів дисертаційного дослідження у роботі сільськогосподарських підприємств, застосування даної системи для підтримки прийняття рішень в управлінні аграрним сектором громади, району та області.

Даний документ не є підставою для виникнення будь-яких, в тому числі фінансових зобов'язань.

Директор
ТОВ НВП «Агроресурссистеми»



Олена МЕДВЕДЕНКО

АКТ
впровадження результатів дисертаційного дослідження
Тетяни Іванівни Просянкіної-Жарової на тему:
«Інформаційні технології для систем підтримки прийняття рішень в
управлінні регіональним розвитком»

Результати дисертаційного дослідження Т. І. Просянкіної-Жарової розглянуті фахівцями Департаменту. Встановлено, що запропоновані інформаційні технології розширюють можливості системи соціальної і професійної адаптації за рахунок розширення аналітичної складової використовуваних інформаційних технологій, основу якої становлять математичні моделі, методи інтелектуального аналізу даних, текстова аналітика, інформаційні технології оброблення великих масивів структурованих та неструктурованих даних. А саме, застосування пропонованих в роботі моделей, методів, інформаційних технологій дозволить підвищити якість надання консультативних, інформаційних, профорієнтаційних послуг в рамках реалізації програм соціальної та професійної адаптації регіональні особливості ринку праці, потреби різних категорій населення у соціальних послугах, передбачити зміну структури контингенту осіб, що потребують соціального захисту та соціального забезпечення.

Важливими є й пропозиції щодо застосування текстової аналітики та штучного інтелекту у роботі із зверненнями громадян. Використання методики оброблення електронних звернень та інформації з соціальних мереж дозволить не лише скоротити строки опрацювання звернень, а й виявляти потреби різних верств населення у соціальному захисті та соціальному забезпеченні, аналізувати рівень задоволення існуючими соціальними програмами.

Отже, вважаємо, що результати дисертаційного дослідження можуть бути застосовані в роботі регіональних управлінь Соціального захисту населення та адаптовані до використання у Єдиній інформаційній системі соціальної сфери України.

Даний документ не є підставою для виникнення будь-яких, в тому числі фінансових зобов'язань.

Т.в.о. директора департаменту



Тетяна КРУТ



**ПАЛАНСЬКА СІЛЬСЬКА РАДА
УМАНСЬКОГО РАЙОНУ ЧЕРКАСЬКОЇ ОБЛАСТІ
ВИКОНАВЧИЙ КОМІТЕТ**

вулиця Грушевського, 11, с. Паланка, Уманський район, Черкаська область, 20340
тел. (04744) 9-94-31 e-mail: palanka-otg@ukr.net код ЄДРПОУ 04409878

Акт

**про впровадження результатів дисертаційного дослідження
Просянкиної-Жарової Тетяни Іванівни
на тему «Інформаційні технології для систем підтримки прийняття
рішень в управлінні регіональним розвитком»**

Результати даного дисертаційного дослідження, використані у роботі Паланської сільської ради та її виконавчого комітету у процесі розроблення пропозицій щодо цифровізації роботи громади (інформаційна технологія роботи із зверненнями громадян), а також у процесі підготовки та прийняття рішень про затвердження плану заходів щодо складання прогнозу бюджету громади на 2026-2028 роки, плану заходів щодо складання проекту бюджету Паланської сільської територіальної громади на 2026 рік та підготовці пропозицій щодо Програми соціально-економічного розвитку Паланської сільської територіальної громади на 2026-2028 роки, інвестиційного забезпечення її агропромислового комплексу (моделі та методи аналізу та прогнозування розвитку соціально-еколого-економічних систем в умовах невизначеності).

Даний документ не є підставою для проведення будь-яких розрахунків та передачі майнових чи авторських прав, виникнення фінансових та інших зобов'язань.

Керуюча справами
(секретар) виконавчого комітету



ШВЕЦЬ Валентина ШВЕЦЬ

Акт
про впровадження результатів дисертаційного дослідження
Присянкіної-Жарової Тетяни Іванівни
на тему «Інформаційні технології для систем підтримки прийняття рішень в
управлінні регіональним розвитком»

Результати даного дисертаційного дослідження використані у роботі Уманської міської ради при розробленні та реалізації Міської цільової програми інформатизації м. Умань на 2021-2025 роки та бюджетів Уманської міської територіальної громади на 2024 та 2025 роки.

На увагу заслуговує запропонована в роботі інформаційна технологія підтримки прийняття рішень, основу якої становлять методи інтелектуального аналізу даних, сценарний аналіз, математичні, економетричні, когнітивні моделі та ансамблі моделей.

Вхідною інформацією для побудови моделей є статистичні дані, показники звітності територіальної громади, наявні у відкритому доступі. Математичні моделі, пропонувані в роботі можуть бути використані при формуванні бюджетних запитів, прогнозуванні показників бюджету, визначення спроможності громади, перспектив її розвитку, формуванні інвестиційних паспортів громад. Результати прогнозування можна легко «прив'язати» до геоданих завдяки використанню процедури побудови діаграми розсіювання. Застосування даного підходу дозволяє підвищити якість та обґрунтованість управлінських рішень у системі місцевого самоврядування, зокрема, у процесі розроблення місцевого бюджету, цільових, інвестиційних програм, тощо в умовах невизначеності та ризику.

Корисною є й представлена в роботі методика аналізу звернень громадян та дослідження суспільної думки, особливістю якої є використання засобів text mining. Впровадження даної методики у роботу органів місцевого самоврядування дозволить підвищити оперативність вирішення порушених у зверненнях питань, отримати об'єктивну оцінку роботи органів місцевого самоврядування, окреслити потреби громади у соціальному захисті та соціальному забезпеченні, освітніх та медичних послугах.

За нашими оцінками представлена інформаційна технологія має наукову новизну та практичну цінність для територіальних громад різного типу. Результати апробації інформаційної технології як в цілому, так і окремих моделей та методів аналізу, обробки геопросторових даних, призначені для прогнозування показників розвитку територіальних громад, дозволяють зробити висновки щодо перспективності їх використання, як з метою оптимізації галузевої структури економіки територіальної громади, так і підвищення її інвестиційної привабливості, визначення шляхів задоволення соціальних потреб населення.

Даний документ не є підставою для проведення будь-яких розрахунків та передачі майнових чи авторських прав, виникнення будь-яких, в тому числі фінансових зобов'язань.

Секретар виконавчого комітету
Уманської міської ради



Наталія СИРОТЮК



ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ «ПС СОЛЮШЕНЗ»

Код ЄДРПОУ 43123101

Юридична адреса: 01103 м. Київ, вул. Михайла Бойчука, буд. 12
тел. (044) 22-33-502

05.08.2025

АКТ

впровадження результатів дисертаційного дослідження

Тетяни Іванівни Просянкиної-Жарової на тему:

«Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком»

Місцеві бюджети є ключовим інструментом регіонального розвитку, оскільки саме через них забезпечується фінансування соціальних програм, інфраструктурних проєктів та інших важливих ініціатив місцевого самоврядування. В умовах децентралізації та зростання фінансової автономії територій особливого значення набуває питання ефективного управління бюджетними ресурсами. Тому вдосконалення інформаційно-аналітичного забезпечення бюджетного процесу, розробка сучасних інструментів підтримки прийняття рішень стають критично важливими не лише з теоретичної, а й з практичної точки зору.

Сучасні інформаційні технології, представлені в дослідженні, спрямовані на оптимізацію аналітичної діяльності органів державного та публічного управління. Вони дозволяють покращити комунікацію між учасниками бюджетного процесу, забезпечуючи прозорість, оперативність обміну даними та зручність їх аналізу. Завдяки поєднанню математичного моделювання, економетричного аналізу, інтелектуального аналізу великих масивів даних (Big Data), сценарного прогнозування та рекомендаційних систем, ці рішення дають можливість підвищити точність бюджетного планування, якість моніторингу виконання бюджетів та обґрунтованість управлінських рішень.

Результати дисертаційного дослідження знайшли практичне застосування у роботі ТОВ «ПС Солюшенз» під час випробування програмного комплексу ГІАС «Фінансові ресурси». Ця система призначена для використання органами місцевого самоврядування у плановій, звітній та аналітичній діяльності. Вона спрощує процес формування бюджету, дозволяє оцінювати фінансові ризики, моделювати різні сценарії розвитку подій та отримувати рекомендації щодо оптимального розподілу коштів.

Важливо зазначити, що даний документ має інформаційний характер і не є підставою для виникнення фінансових зобов'язань. Його мета – продемонструвати потенціал сучасних інформаційних технологій у сфері управління місцевими фінансами та підкреслити їх роль у забезпеченні сталого розвитку регіонів.

З повагою,
Директор ТОВ «ПСС»



Орленко І.М.



Algoritm-X

ТОВ «АЛГОРИТМ-Х»
ЄДРПОУ 45151557
вул. Вікторія Хвойки, 18/14
м. Київ, Україна, 04080
тел.: +38(044) 221-74-82
e-mail: info@algoritm-x.ua

№ _____

АКТ

впровадження результатів дисертаційного дослідження
Тетяни Іванівни Просянкіної-Жарової на тему:
«Інформаційні технології для систем підтримки прийняття рішень в
управлінні регіональним розвитком»

Місцеві бюджети відіграють важливу роль у регіональному розвитку, забезпечуючи фінансовими ресурсами реалізацію повноважень місцевого самоврядування. Тому питання удосконалення інформаційно-аналітичного забезпечення бюджетного процесу, підтримки прийняття рішень в управлінні регіональним розвитком має не лише теоретичне, а й практичне значення. Представлені в роботі інформаційні технології призначені для використання у аналітичній діяльності органів державного та публічного управління забезпечують поліпшення взаємодії та обміну інформацією між учасниками бюджетного процесу, дозволяють підвищити якість обґрунтування, планування, моніторингу виконання місцевих бюджетів за рахунок поєднання математичного та економетричного моделювання, інтелектуального аналізу даних, сценарного аналізу та рекомендаційних систем. Тому, результати дисертаційного дослідження використані у роботі ТОВ «АЛГОРИТМ-Х» при розробленні програмного забезпечення ГІАС «Фінансові ресурси», призначеного для використання у плановій та аналітичній роботі органів місцевого самоврядування.

Даний документ не є підставою для виникнення фінансових зобов'язань.

Директор ТОВ «АЛГОРИТМ-Х»



 Богдан ЖОВТОБРЮХ

ТОВ «СП «УКРІНТЕРМ»»



"JV 'UKRINTERM'" LLC

вул. Івана Калитишина, 107-А, м. Київ (Україна)
Київська обл., 09114, Україна
E-mail: info@ukrinterm.com.ua
«AT "РАЙФФАЙЗЕН БАНК"» м. Київ МФО 300003
ЄДРПОУ: 20016760, ДНУ 20016760278
Єдиного 30 200172604
Тел.: (044) 513-99-1, 19-111-2
Факс: (044) 59-73-23, 19-111-2
e-mail: info@ukrinterm.com.ua, www.ukrinterm.com.ua

ISO 9001
ISO 14001

107-A, Ivana Kalitishyna str., Kyiv (Ukraine)
Kyiv region, 09114, Ukraine
E-mail: info@ukrinterm.com.ua
JSC "RAIFFAISEN BANK", MFO 300003
EDRPOU: 20016760, DNU 20016760278
EAT Certificate 200172604
Tel.: (+380-4965) 9-111-2, 19-98-1
Fax: (+380-4965) 9-73-23, 8-111-2
e-mail: info@ukrinterm.com.ua, www.ukrinterm.com.ua

АКТ

про впровадження результатів дисертаційного дослідження Просвякної-Жарової Тетяни Іванівни на тему «Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком»

засвідчує те, що результати даного дисертаційного дослідження використані у роботі ТОВ «СП «Укрінтерм»» й мають практичну цінність. Адже втілення ефективного енергоменеджменту – важлива складова регіонального розвитку: чинник сталого економічного зростання регіону і громад та забезпечення належних умов для проживання населення. Тому запропонована методика прогнозування навантаження енергосистеми, яка поєднує сценарний аналіз та математичне моделювання, дозволяє отримати пропозиції високої якості як на довгостроковий, середньостроковий, так і короткостроковий період. Представлений у роботі підхід доповнює існуючі методики розроблення регіональних програм енергозбереження. Його застосування є особливо корисним на етапі прийняття рішень, оскільки дозволяє розглядати декілька сценаріїв розвитку подій, враховувати вплив різних груп чинників – як технологічних, технічних, так і економічних та екологічних, тощо.

Отже, вважаємо, що результати дисертаційного дослідження «Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком», мають практичну цінність, можуть бути використані у процесі прийняття рішень в управлінні регіональним розвитком, зокрема, розробленні регіональних ініціатив та програм енергозбереження.

Даний документ не є підставою для виникнення будь-яких, в тому числі фінансових зобов'язань.

Генеральний директор
ТОВ «СП «Укрінтерм»»

Мороз П.М.



Громадська спілка
"Асоціація інженерів сталих
енергетичних технологій України"





УКРАЇНА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

(КПІ ім. Ігоря Сікорського)

пр-т Берестейський, 37, м. Київ, 03056, тел. (044) 204 82 82 тел. (044) 204 94 94
<https://www.kpi.ua> e-mail: mail@kpi.ua ЄДРПОУ 02070921

22.07.2025 № 13/09.00
на № _____ від _____

АКТ

**про впровадження результатів дисертаційного дослідження
Присянкіної-Жарової Тетяни Іванівни
на тему «Інформаційні технології для систем підтримки прийняття рішень
в управлінні регіональним розвитком» у освітній процес**

Результати дисертаційного дослідження на тему «Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком» використані у освітньому процесі кафедри математичних методів системного аналізу при викладанні дисциплін «Фундаментальний аналіз фінансового ринку» та «Системи і методи підтримки прийняття рішень» для студентів другого (магістерського) освітнього рівня спеціальності 124 Системний аналіз (освітні програми «Системний аналіз і управління» та «Системний аналіз фінансового ринку») та дисципліни «Мікро- та макроекономічні системи» для здобувачів вищої освіти першого (бакалаврського) рівня, спеціальності 122 Комп'ютерні науки (освітньо-професійна програма «Системи і методи штучного інтелекту»), 124 Системний аналіз (освітньо-професійна програма «Системний аналіз і управління»).

В.о. директора ІНПСА

Завідувач кафедри математичних
методів системного аналізу

Костянтин СФРЕМОВ

Оксана ТИМОЩУК



ЗАТВЕРДЖУЮ

Васильковецький сільський голова

Гайдаш А. В.

" " липня 2025 р.

АКТ

впровадження наукових результатів дисертації
на тему «Інформаційні технології для систем підтримки прийняття рішень в
управлінні регіональним розвитком»
докторанта Інституту телекомунікацій і глобального інформаційного простору
Національної академії наук України
ПРОСЯНКІНОЇ-ЖАРОВОЇ Тетяни Іванівни

Комісія Васильковецької територіальної громади Тернопільської області у складі голови комісії, начальника фінансового відділу Безкоровайної Галини Михайлівної та членів комісії: завідувача сектору економіки, інфраструктури та житлово-комунального господарства Хортика Богдана Степановича, директора комунальної установи "Центр надання соціальних послуг" Кіндзер Ганни Іванівної розглянула запропоновані матеріали, які представляють результати дисертаційного дослідження докторантки Просянкіної-Жарової Тетяни Іванівни на тему «Інформаційні технології для систем підтримки прийняття рішень в управлінні регіональним розвитком».

Комісією встановлено, що розв'язання поставленого наукового завдання забезпечує зростання якості інформаційно-аналітичної діяльності у процесі підтримки прийняття рішень в управлінні розвитком територіальної громади шляхом впровадження інформаційних технологій аналізу та прогнозного моделювання соціально-економічних показників, що описують її структуру та динаміку для виявлення шляхів забезпечення сталого розвитку громади та дотримання соціальних стандартів.

Комісією також встановлено, що запропоновані інформаційні технології дозволяють враховувати особливості територіальних громад різних типів, вплив основних груп зовнішніх та внутрішніх чинників на їх розвиток, виконання місцевого бюджету, цільових програм та планів, вчасно реагувати на відхилення, загрози, коригувати управлінські рішення відповідно до поточної ситуації, що забезпечить стійкість територіальних громад, як їх здатність адаптуватись до криз та протистояти зовнішнім шокам.

Практичне значення результатів дисертаційного дослідження полягає в тому, що поєднання запропонованих математичних моделей (та їх ансамблів), комп'ютерного моделювання, методів інтелектуального аналізу даних, сценарного аналізу та когнітивного моделювання у єдиній інформаційній технології, забезпечить високу якість прогнозів соціально-економічних показників, використовуваних при розробленні бюджету громади, планів, цільових програм, інвестиційного паспорту територіальної громади, тощо, що сприяє покращенню соціально-економічного розвитку територіальної громади завдяки підвищенню якості рішень в управлінні її розвитком, що приймаються в умовах невизначеності, спричиненої дією різних чинників та загроз, в тому числі, пов'язаними із військовим конфліктом.

Отже, вважаємо, що результати дисертаційного дослідження можуть бути застосовані в роботі органів місцевого самоуправління.

Голова комісії:

Члени комісії:



Галина БЕЗКОРОВАЙНА

Богдан ХОРТИК

Ганна КИНДЗЕР

Код для завантаження дані щодо захворюваності з Інтернет

/* -----

Вихідний файл:

D: \ Work \ Covid19_data_ \ table_ECDC_CORONA_VIRUS.csv

Сервер: Локальна файлова система

Вихідні дані: WORK.table_ECDC_CORONA_VIRUS

Сервер: Local

----- */

DATA WORK.table_ECDC_CORONA_VIRUS;

LENGTH

date 8

day 8

month 8

year 8

confirmed 8

deaths 8

country_orig \$ 42

country_state_id \$ 8

Country_territory_code \$ 7

popData2019 8

continentExp \$ 7

Cumulative_number_for_14_days_of 8

isoalpha2 \$ 2

country_id \$ 2

country \$ 33

ISO 8

latitude 8

longitude 8

x 8

y 8

confirmed_total 8
 deaths_total 8
 SP_POP_TOTL 8
 EN_POP_DNST 8
 SP_POP_0014_TO 8
 SP_POP_1564_TO 8
 SP_POP_65UP_TO 8
 today_date 8
 confirmed_all 8
 deaths_all 8
 days_since_first_case 8
 days_since_tenth_case 8
 days_since_hundredth_case 8
 days_since_first_death 8
 days_since_tenth_death 8
 days_since_hundredth_death 8 ;

FORMAT

date MMDDYY10.
 day BEST12.
 month BEST12.
 year BEST12.
 confirmed BEST12.
 deaths BEST12.
 country_orig \$CHAR42.
 country_state_id \$CHAR8.
 Country_territory_code \$CHAR7.
 popData2019 BEST12.
 continentExp \$CHAR7.
 Cumulative_number_for_14_days_of BEST12.
 isoalpha2 \$CHAR2.
 country_id \$CHAR2.
 country \$CHAR33.

ISO BEST3.
 latitude BEST12.
 longitude BEST12.
 x BEST12.
 y BEST12.
 confirmed_total BEST12.
 deaths_total BEST12.
 SP_POP_TOTL BEST12.
 EN_POP_DNST BEST12.
 SP_POP_0014_TO BEST12.
 SP_POP_1564_TO BEST12.
 SP_POP_65UP_TO BEST12.
 today_date DDMMYY10.
 confirmed_all BEST12.
 deaths_all BEST12.
 days_since_first_case BEST12.
 days_since_tenth_case BEST12.
 days_since_hundredth_case BEST12.
 days_since_first_death BEST12.
 days_since_tenth_death BEST12.
 days_since_hundredth_death BEST12. ;
 INFORMAT
 date MMDDYY10.
 day BEST12.
 month BEST12.
 year BEST12.
 confirmed BEST12.
 deaths BEST12.
 country_orig \$CHAR42.
 country_state_id \$CHAR8.
 Country_territory_code \$CHAR7.
 popData2019 BEST12.

continentExp \$CHAR7.

Cumulative_number_for_14_days_of BEST12.

isoalpha2 \$CHAR2.

country_id \$CHAR2.

country \$CHAR33.

ISO BEST3.

latitude BEST12.

longitude BEST12.

x BEST12.

y BEST12.

confirmed_total BEST12.

deaths_total BEST12.

SP_POP_TOTL BEST12.

EN_POP_DNST BEST12.

SP_POP_0014_TO BEST12.

SP_POP_1564_TO BEST12.

SP_POP_65UP_TO BEST12.

today_date DDMMYY10.

confirmed_all BEST12.

deaths_all BEST12.

days_since_first_case BEST12.

days_since_tenth_case BEST12.

days_since_hundredth_case BEST12.

days_since_first_death BEST12.

days_since_tenth_death BEST12.

days_since_hundredth_death BEST12. ;

INFILE 'C:\Users\SASAdm\AppData\Roaming\SAS\EnterpriseGuide\EGTEMP\SEG-968-368fe2ce\contents\table_ECDC_CORONA_VIRUS-e08deda9a475436fbd25923aedc9e4a3.txt'

LRECL=451

ENCODING="UTF-8"

TERMSTR=CRLF

DLM='7F'x

MISSOVER

DSD ;

INPUT

date : MMDDYY10.

day : BEST12.

month : BEST12.

year : BEST12.

confirmed : BEST12.

deaths : BEST12.

country_orig : \$CHAR42.

country_state_id : \$CHAR8.

Country_territory_code : \$CHAR7.

popData2019 : BEST12.

continentExp : \$CHAR7.

Cumulative_number_for_14_days_of : COMMAX12.

isoalpha2 : \$CHAR2.

country_id : \$CHAR2.

country : \$CHAR33.

ISO : BEST3.

latitude : COMMAX12.

longitude : COMMAX12.

x : COMMAX12.

y : COMMAX12.

confirmed_total : BEST12.

deaths_total : BEST12.

SP_POP_TOTL : BEST12.

EN_POP_DNST : COMMAX12.

SP_POP_0014_TO : BEST12.

SP_POP_1564_TO : BEST12.

SP_POP_65UP_TO : BEST12.

today_date : DDMMYY10.

confirmed_all : BEST12.

deaths_all : BEST12.

```

days_since_first_case : BEST12.
days_since_tenth_case : BEST12.
days_since_hundredth_case : BEST12.
days_since_first_death : BEST12.
days_since_tenth_death : BEST12.
days_since_hundredth_death : BEST12. ;
RUN;

```

Додаток Б

Код програми для обчислення кумулятивних статистик

```

%let file_in = WORK.TABLE_ECDC_CORONA_VIRUS;
%let country=UKRAINE;
%symdel country;
%MACRO ETL(country);
data temp;
format date Date9.;
set &file_in. (where=(UPCASE(country_orig)="&country."));
run;
data temp2;
format date Date9.;
set &file_in. (where=(UPCASE(country_orig)="&country."));
tot_confirmed+confirmed;
tot_deaths+deaths;
if tot_confirmed>=100 then Start_Date+1;
week=int(Start_Date/8)+1;
IF Start_Date=0 then Delete;
run;
proc sort data=temp2 out=temp2_sort;
by week;
run;
data &country;

```

```
set temp2_sort;
by week;
if First.week=1 then output;
Keep country week tot_confirmed tot_deaths;
run;
%MEND ETL;
%ETL(UKRAINE);
%ETL(BELARUS);
%ETL(HUNGARY);
%ETL(NIGERIA);
%ETL(ROMANIA);
%ETL(SERBIA);
```

```
data virus;
set
WORK.UKRAINE
WORK.BELARUS
WORK.HUNGARY
WORK.NIGERIA
WORK.ROMANIA
WORK.SERBIA;
run;
```

Додаток В

Код програми для дослідження подібності процесів

```
libname fdv 'C:\Work\DataDaseVirus';
```

```
DATA FDV.confirmed;
```

```
INFILE DATALINES DLM=' ' MISSOVER DSD ;
```

```
INPUT
```

```
Week      : BEST32.
```

```
Ukraine   : BEST32.
```

```
Belarus   : BEST32.
```

```
Hungary   : BEST32.
```

```
Nigeria   : BEST32.
```

```
Romania   : BEST32.
```

```
Serbia    : BEST32. ;
```

```
DATALINES;
```

```
1 113 105 103 131 113 114 126
```

```
2 794 700 343 238 367 495 384
```

```
3 1892 3281 733 373 1760 2777 1476
```

```
4 4462 7281 1458 873 4057 10131 3380
```

```
5 8125 14027 2098 1932 6879 32008 6318
```

```
6 11913 21101 2727 3912 9710 74588 8275
```

```
7 15648 28681 3150 5621 12240 134687 9677
```

```
8 18876 36198 3417 7839 14811 221344 10295
```

```
9 21905 43403 3713 10578 16704 299941 10919
```

```
10 25411 50265 3867 13464 18070 370680 11354
```

```
11 29753 56032 3970 17735 19398 441108 11741
```

```
12 35825 60382 4077 22614 20749 511423 12310
```

```
13 42982 62997 4107 27564 22760 576952 12990
```

```
14 49043 64767 4157 31987 25286 634437 14564
```

```
15 54771 66095 4220 36663 28166 687862 17076
```

```
16 60995 67251 4293 41180 32079 739947 19717
```

```
17 68794 68250 4435 44433 37458 789190 22852
```

```

18 78261 69102 4535 47743 45902 834499 25882
19 89719 69950 4731 50964 55241 877135 28099
20 104958 71346 4970 53477 65177 917884 29782
21 121215 72663 5288 54905 74963 956749 30820
22 143030 . 6923 . 84468 995319 31581
23 . . . . 93864 1035789 .

```

```
;
```

```
run;
```

```
ods graphics on;
```

```

proc similarity data=fdv.CONFIRMED out=_null_ outpath=path
print=all plot=all;
input Ukraine;
target Belarus / measure=absdev slide=none
compress=(localabs=2)
expand=(localabs=2);
run;

```

```

proc similarity data=fdv.CONFIRMED out=_null_ outpath=path
print=all plot=all;
input Ukraine;
target Hungary / measure=absdev slide=none
compress=(localabs=2)
expand=(localabs=2);
run;

```

```

proc similarity data=fdv.CONFIRMED out=_null_ outpath=path
print=all plot=all;
input Ukraine;
target Nigeria / measure=absdev slide=none
compress=(localabs=2)

```

```
expand=(localabs=2);
```

```
run;
```

```
proc similarity data=fdv.CONFIRMED out=_null_ outpath=path
```

```
print=all plot=all;
```

```
input Ukraine;
```

```
target Romania / measure=absdev slide=none
```

```
compress=(localabs=2)
```

```
expand=(localabs=2);
```

```
run;
```

```
proc similarity data=fdv.CONFIRMED out=_null_ outpath=path
```

```
print=all plot=all;
```

```
input Ukraine;
```

```
target Serbia / measure=absdev slide=none
```

```
compress=(localabs=2)
```

```
expand=(localabs=2);
```

```
run;
```

Додаток Д

Код програми для реалізації обчислювального алгоритму навчання ШНМ

Listing of the module C_ANN_training.py

Script 1. Initialization of the ANN parameters

Python programming language script for initialization of the neuron network parameters

Line number	The program code
1	import sys, openpyxl
2	import numpy as np
3	from os.path import join, abspath

```

4
5     # defining a function to calculate the standard logistic activation function
6     def logistic(x):
7         return 1.0 / (1 + np.exp(-x))
8
9     # defining a function to calculate the derivative of the activation function
10    def logistic_deriv(x):
11        return logistic(x) * (1 - logistic(x))
12
13    # network learning parameters
14    epoch_count = 100 # learning epochs
15    alpha = 0.002 # the learning speed coefficient value
16
17    # reading network parameters and training examples
18    # from MS Excel data file
19    # open the specified file only to read the values
20    # (not read formulas)
21
22    # it is created a full path to open the file based on
23    # the directory in which the program is running
24    # such full name is stored in data_path
25    data_path = join('.', 'Data', "Training_Data.xlsx")
26    data_path = abspath(data_path)
27
28    Training_Data = openpyxl.open(data_path, read_only=True, data_only=True)
29
30    # select the first worksheet in the MS Excel filebook
31    sheet_data = Training_Data.worksheets[0]
32
33    max_row_data = sheet_data['A3'].value # number of teaching examples
34    # read directly from the file (number of input-output network pairs)
35

```

```

36 # network parameters
37 input_size = sheet_data['A5'].value # number of inputs
38 hidden_size = sheet_data['A7'].value # number of neurons in the hidden layer
39 output_size = sheet_data['A9'].value # number of outputs
40
41 print('the ANN parameters ')
42 print('number of inputs:', input_size)
43 print('number of neurons in the hidden layer:', hidden_size)
44 print('number of outputs:', output_size)
45 print('number of teaching examples:', max_row_data)
46
47 # definition and initialization of matrices of the ANN input and output values
48 len_ryadok = input_size
49
50 characteristics_riverbed = np.zeros(
51     shape=(max_row_data, len_ryadok))
52 # input - hydromorphological channel characteristics B, h, Sf, n
53 coef_C = np.zeros(shape=(max_row_data)) # output – the Chézy coefficient value
54
55 # filling the valid matrices of input and output values
56 # of the network with training data from the file
57
58 for i in range(max_row_data):
59     coef_C[i] = sheet_data[i + 2][input_size + 1].value
60     for j in range(len_ryadok):
61         characteristics_riverbed[i][j] = sheet_data[i + 2][j + 1].value
62
63 # matrix of weight coefficients W_1 i W_2
64 # are given by random values
65
66 np.random.seed(1)
67 W_1 = 0.02 * np.random.random((input_size, hidden_size)) - 0.01

```

68	<code>W_2 = 0.6 * np.random.random((hidden_size, output_size)) - 0.3</code>
69	

Script 2. Direct and reverse courses of calculations, and checking the ANN training adequacy

Python programming language script for direct and reverse courses of calculations,
and checking the ANN training adequacy

Line number	The program code
70	<code># realization of learning epochs</code>
71	<code># (alternate execution of direct and reverse)</code>
72	<code># for all learning examples for 1 epoch</code>
73	
74	<code>correct = 0 # number of satisfactory learning outcomes</code>
75	<code>total = 0 # number of studied learning examples</code>
76	<code>error_delta = 0.001 # the network error value</code>
77	
78	<code>for iteration in range(epoch_count):</code>
79	<code> layer_2_e2 = 0 # the sum of squares deviations</code>
80	<code> for i in range(len(characteristics_riverbed)):</code>
81	<code> # direct course</code>
82	<code> layer_0 = characteristics_riverbed[i:i + 1]</code>
83	<code> layer_1 = logistic(np.dot(layer_0, W_1))</code>
84	<code> layer_2 = np.dot(layer_1, W_2)</code>
85	
86	<code> layer_2_e2 += np.sum((layer_2 - coef_C[i:i + 1]) ** 2)</code>
87	
88	<code> # reverse course</code>
89	<code> layer_2_delta = (coef_C[i:i + 1] - layer_2)</code>
90	<code> layer_1_delta = layer_2_delta.dot(W_2.T) * logistic_deriv(layer_1)</code>
91	
92	<code> W_2 = W_2 + alpha * layer_1.T.dot(layer_2_delta)</code>
93	<code> W_1 = W_1 + alpha * layer_0.T.dot(layer_1_delta)</code>

```

94
95     # training accuracy assessment
96     if (np.abs(layer_2_delta) < error_delta):
97         correct += 1
98     total += 1
99     if ((iteration == epoch_count - 1) and (i % 10 == 9)):
100         print("Iteration:" + str(iteration)+", step:"+str(i))
101         print("_C_ = " + str(layer_2))
102         print("C = " + str(coef_C[i]))
103         print("Quadratic Error:" + str(layer_2_e2))
104         print("Training Accuracy:" + str(correct * 100 / float(total)))
105
106     # checking the accuracy (adequacy) of forecasting
107     # reading and downloading a test sample of data
108
109     # selecting the second worksheet with test examples in the MS Excel file book
110     sheet_data = Training_Data.worksheets[1]
111
112     # the number of test examples is read directly from the file
113     # (number of input-output network pairs)
114     max_row_test = sheet_data['A3'].value
115
116     characteristics_riverbed_test = np.zeros(
117         shape=(max_row_test, len_ryadok))
118     # input - test values of hydromorphological channel characteristics B, h, Sf, n
119     coef_C_test = np.zeros(shape=(max_row_test))
120     # output - the Chézy coefficient test values
121
122     # filling the network valid matrices of input and output values
123     # with test examples
124     for i in range(max_row_test):
125         coef_C_test[i] = sheet_data[i + 2][input_size + 1].value

```

126	for j in range(len_ryadok):
127	characteristics_riverbed_test[i][j] = sheet_data[i + 2][j + 1].value
128	
129	correct = 0 # number of satisfactory forecasting results
130	total = 0 # number of checked test examples
131	
132	# calculation of network outputs and accuracy assessment
133	for i in range(max_row_test):
134	#direct course of calculations
135	layer_0 = characteristics_riverbed_test[i:i+1]
136	layer_1 = logistic(np.dot(layer_0,W_1))
137	layer_2 = np.dot(layer_1,W_2)
138	
139	# forecasting accuracy assessment
140	if(np.abs(coef_C_test[i:i+1]-layer_2) < error_delta):
141	correct += 1
142	total += 1
143	print('example ', i, ', test C: ', coef_C_test[i], ', calculated C: ', layer_2)
144	print('all test examples: ', total)
145	print("Test Accuracy:" + str(correct*100 / float(total)))
146	# forecasting accuracy on test examples
147	
148	Training_Data.close() # closing of the MS Excel data file
149	

Script 3. Storage of the ANN learning outcomes

Python programming language script for Storage of the ANN learning outcomes

Line number	The program code
150	# entry in the text file of learning outcomes - matrices of weights W_1 та W_2
151	##(if the file contains information, data will be deleted)
152	

```

153 # it is created a full path to open the file based on
154 # the directory in which the program is running
155 # such full name is stored in data_path
156 data_path = join('.', 'Data', "weights_matrix_1.txt")
157 data_path = abspath(data_path)
158
159 f = open(data_path, 'w')
160 for i in range (0,input_size-1):
161     for j in range(0,hidden_size-1):
162         f.write(str(W_1[i][j])+ ' ')
163 # writing to the file only the values of the array through space, others are not transmitted
164     f.write(str(W_1[i][hidden_size-1])+'\n')
165 for j in range(0,hidden_size-1):
166     f.write(str(W_1[input_size-1][j])+ ' ')
167 f.write(str(W_1[input_size-1][hidden_size-1]))
168 f.close()
169
170 # it is created a full path to open the file based on
171 # the directory in which the program is running
172 # such full name is stored in data_path
173 data_path = join('.', 'Data', "weights_matrix_2.txt")
174 data_path = abspath(data_path)
175
176 f = open(data_path, 'w')
177 for i in range (0,hidden_size-1):
178     for j in range(0,1):
179         f.write(str(W_2[i][j])+'\n')
180 f.write(str(W_2[hidden_size-1][0]))
181 f.close()
182
183 print (' The adjusted matrices of weight coefficients successfully saved.')
184

```

185	<code>print('Press any key to end the program.')</code>
186	<code>input()</code>

Додаток Д

Listing of the module C_ANN_calculating.py

Python programming language script for forecasting the Chézy roughness coefficient

Line number	The program code
1	<code>import openpyxl</code>
2	<code>import numpy as np</code>
3	<code>from openpyxl import Workbook</code>
4	<code>from openpyxl.styles import Alignment, PatternFill, Font</code>
5	<code>from os.path import join, abspath</code>
6	
7	<code># calculating the Chézy roughness coefficient C value</code>
8	<code># with using the trained ANN (adjusted matrices of weights),</code>
9	<code># weight matrices are downloaded from text files,</code>
10	<code># input arguments are loaded from a MS Excel file,</code>
11	<code># the calculation results are stored in a MS Excel file</code>
12	
13	<code># defining a function to calculate the derivative of the activation function</code>
14	
15	<code>def logistic(x):</code>
16	<code> return 1.0 / (1 + np.exp(-x))</code>
17	
18	<code># open the specified file only to read the values</code>
19	<code>data_path = join('.', "Input.xlsx")</code>
20	<code>data_path = abspath(data_path) # it is created a full path to open the file based on</code>
21	
22	<code># the directory in which the program is running</code>

```

23 Input_Data = openpyxl.open(data_path, read_only=True, data_only=True)
24 sheet_data = Input_Data.worksheets[0]
25
26 max_row_data = sheet_data['A3'].value # number of input data examples
27 # to calculate the Chézy roughness coefficient C value
28
29 # network parameters corresponding to the input data
30 input_size = sheet_data['A5'].value # number of inputs
31 hidden_size = sheet_data['A7'].value # number of neurons in the hidden layer
32 output_size = sheet_data['A9'].value # number of outputs
33
34 # definition and initialization of matrices of the ANN input values and
35 # matrix of the Chézy roughness coefficient calculated values
36
37 len_ryadok = input_size
38 characteristics_riverbed = np.zeros(
39     shape=(max_row_data, len_ryadok))
40 # input - hydromorphological channel characteristics
41
42 coef_C = np.zeros(shape=(max_row_data)) # output – the Chézy coefficient value
43
44 # loading input values
45 for i in range(max_row_data):
46     for j in range(len_ryadok):
47         characteristics_riverbed[i][j] = sheet_data[i + 2][j + 1].value
48
49 print('The input data for the ANN was downloaded from the file:')
50 print(data_path)
51 print('the ANN parameters:')
52 print('number of inputs -', input_size)
53 print('number of neurons in the hidden layer -', hidden_size)
54 print('number of outputs -', output_size)

```

```

55     print('Number of input examples:', max_row_data)
56
57     list_param=[] # list of descriptions of input parameters
58     print('The coefficient C/100 is being investigated', end=' ')
59     for j in range(len_ryadok):
60         list_param.append(sheet_data[1][j+1].value)
61         print(sheet_data[1][j+1].value, end=',')
62     print()
63
64     Input_Data.close() # closing of the MS Excel input data file
65
66     # reading of matrix of weight coefficients W_1 i W_2
67
68     data_path = join('.', 'Data', "weights_matrix_1.txt")
69     data_path = abspath(data_path)
70
71     f = open(data_path)
72     raw_matrix1 = f.readlines()
73     f.close()
74
75     data_path = join('.', 'Data', "weights_matrix_2.txt")
76     data_path = abspath(data_path)
77
78     f = open(data_path)
79     raw_matrix2 = f.readlines()
80     f.close()
81
82     # initialization of matrices W_1 i W_2
83     W_1 = np.zeros(
84         shape=(input_size, hidden_size))
85     W_2 = np.zeros(shape=(hidden_size, output_size))
86

```

```

87      # filling of matrices W_1 i W_2
88
89      for i in range (len(raw_matrix1)):
90          b=raw_matrix1[i]
91          a=b.split(' ')
92          # len(a) == len(raw_matrix2) - number of the matrix rows W_2 = columns W_1
93          for j in range(len(raw_matrix2)):
94              W_1[i][j]=float(a[j])
95
96      for i in range (len(raw_matrix2)):
97          W_2[i]=float(raw_matrix2[i])
98
99      if (len(raw_matrix1)==input_size)and(len(raw_matrix2)==hidden_size):
100          print(' the input data correspond to the parameters of the ANN matrices of weights,')
101          # calculation of the network outputs
102          for i in range(max_row_data):
103              # direct course of calculations
104              layer_0 = characteristics_riverbed[i:i + 1]
105              layer_1 = logistic(np.dot(layer_0, W_1))
106              layer_2 = np.dot(layer_1, W_2)
107              coef_C[i] = layer_2
108              print('for a set of parameters №',i+1,', it is calculated C/100 =', coef_C[i])
109
110      # creating a MS Excel file to write data to a spreadsheet
111      Out_Data=Workbook()
112      sheet_data=Out_Data.active
113      sheet_data.title="output_data"
114
115      # recording data in the table
116      # column names are added
117      sheet_data['A1'].value='j'
118      sheet_data.cell(row=1, column=len_ryadok+2).value='C/100'

```

```

119     for j in range(len_ryadok):
120         sheet_data[1][j + 1].value=list_param[j]
121
122     # recording the input data and the corresponding calculated values of C
123     for i in range(max_row_data):
124         sheet_data['A'+str(i+2)].value = i+1
125         sheet_data.cell(row=i+2, column=len_ryadok + 2).value = coef_C[i]
126         for j in range(len_ryadok):
127             sheet_data[i + 2][j + 1].value = characteristics_riverbed[i][j]
128
129     # formatting the output to the table
130     for i in range(1, len_ryadok+2):
131         zag = sheet_data.cell(row=1, column=i)
132         zag.alignment = Alignment(horizontal='center')
133         zag.font = Font(bold=True, italic=False, color='DC143C', size=12)
134
135         zag = sheet_data.cell(row=1, column=len_ryadok + 2)
136         zag.alignment = Alignment(horizontal='center')
137         zag.font = Font(bold=True, italic=False, color='000000', size=12)
138
139     for i in range(max_row_data+1):
140         sheet_data.cell(row=i+1, column=len_ryadok+2).fill=PatternFill(
141             fill_type='solid', start_color='90EE90', end_color='90EE90')
142
143     # creating (join from parts) file name for storage
144     exfilename = join('.', ('Output.xlsx'))
145     # creating a complete path to store the file
146     exfilename=abspath(exfilename)
147
148     Out_Data.save(exfilename) # storing the file
149     Out_Data.close()
150     print("Saving the calculation results in the file: ")

```

151	print(exfilename)
152	
153	else:
154	print('The input data do not correspond to the ANN matrices of weights.')
155	print('Check that the number of inputs and neurons in the hidden layer match ')
156	print('to the parameters of the trained network in the input file:')
157	print('the number of columns W_1 (the number of neurons in the hidden layer of the
158	network) =', len(raw_matrix2))
159	print('the number of rows W_1 (the number of the ANN inputs) =', len(raw_matrix1))
160	print('the number of rows W_2 (the number of neurons in the hidden layer) =',
161	len(raw_matrix2))
162	
163	input('Press any key to end the program.')
164	