

ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЛІП'ЯНИНА-ГОНЧАРЕНКО ХРИСТИНА ВОЛОДИМИРІВНА



УДК 004.6+004.9

**ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ЗАСАДИ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ
ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ДАНИХ
ТЕРИТОРІАЛЬНИХ ГРОМАД**

05.13.06 – інформаційні технології

РЕФЕРАТ

дисертації на здобуття наукового ступеня
доктора технічних наук

Дисертацією є рукопис.

Робота виконана на кафедрі інформаційно-обчислювальних систем і управління Західноукраїнського національного університету Міністерства освіти і науки України.

Офіційні опоненти:

доктор технічних наук, професор
Безкорвайний Володимир Валентинович,
Харківський національний університет радіоелектроніки;

докторка технічних наук, професорка
Стеценко Інна Вячеславівна,
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»;

доктор технічних наук, професор
Угрин Дмитро Ілліч,
Чернівецький національний університет імені Юрія Федьковича.

Захист відбудеться «14» лютого 2025 р. о 12⁰⁰ годині на засіданні спеціалізованої вченої ради Д70.052.06 в Хмельницькому національному університеті за адресою: 29016, м. Хмельницький, вул. Кам'янецька, 112, ауд. 1-210.

З дисертацією можна ознайомитися в бібліотеці Хмельницького національного університету (29016, м. Хмельницький, вул. Кам'янецька, 110/1).

Реферат оприлюднено «10» січня 2025 р.

Вчений секретар
спеціалізованої вченої ради
к.т.н., доцент



А.О. Нічепорук

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність дослідження. Сучасний розвиток інформаційних технологій та стрімке зростання обсягів соціально-економічних даних, що генеруються територіальними громадами (ТГ), створюють нові можливості та виклики для управління і прийняття рішень. Ці дані є цінним ресурсом для органів місцевого самоврядування, підприємств та інших зацікавлених сторін, оскільки містять інформацію про демографічні зміни, економічну активність, екологічні умови, соціальні настрої тощо. Однак, для того щоби повністю використати потенціал цих даних, необхідно розробити методи їх аналізу, які враховують специфіку нестационарності та високу динамічність соціально-економічних процесів.

Вивченню методів, моделей та засобів інтелектуального аналізу даних присвячено ряд робіт українських та іноземних учених: Саченка А., Бодяньського Є., Бармака О., Говорушенка Т., Гнатчук Є., Комара М., Єгорова Є., Федорової Ю., Чепурненка А., Алєнді А.А., Алама С., Андерссона Дж., Андеобу Л., Арасу Б. С., Бенассі М., Камеро А., Чіанга Л.-Л., Чіу Ч.-Ц., Чивєри М., Коєндадора Б. Е. В., Кости Н. Л., Донга К., Ейхєра А., Ел Морра К., Ел Уаді Дж., Еррусо Х., Еспозіто Е., Етіаті Л., Фаніні Б., Фрєнті П., Фрєйнда Й., Фу Д., Фустоса Дж., Гєдіклі А., Гіті Т.В., Герке С., Гоше А., Гіллєспі А.А., Голдані М. Х., Гомєса М., Гопнараяна А., Говіндасам Р., Грєві А., Гю І.Г.В., Герцоні М., Гуіча Дж. А., Гуо С., Гупті І. П., Гуркана Ф., Гу С. та багатьох інших. У цих роботах висвітлено різні аспекти інтелектуального аналізу даних, зокрема методи машинного навчання, адаптивні алгоритми та їх застосування в різних сферах.

Незважаючи на значний внесок цих досліджень у розвиток методів інтелектуального аналізу даних, існує низка невирішених проблем, пов'язаних зі специфікою соціально-економічних даних ТГ. Зокрема, існуючі методи не завжди враховують нестационарність часових рядів, високу динамічність змін соціально-економічних показників. Це призводить до недостатньої точності прогнозування для прийняття управлінських рішень.

Традиційні підходи до аналізу соціально-економічних даних часто не здатні адаптивно реагувати на швидкі зміни в них, що є критично важливим в умовах сучасних викликів, таких як: економічні кризи, екологічні проблеми, техногенні катастрофи та інформаційні загрози тощо. Відсутність інструментів для аналізу та прогнозування ускладнює процес прийняття рішень, планування розвитку інфраструктури та оптимізації використання ресурсів у ТГ.

Крім того, зростає потреба у впровадженні інтелектуальних систем підтримки прийняття рішень, які б інтегрували передові методи МН, здатні обробляти масиви даних у режимі реального часу. Це особливо актуально для забезпечення соціальної стійкості та розвитку ТГ, управління екологічними ситуаціями, розвитку підприємницької діяльності та просування товарів на інтернет-ринку. Тому існує нагальна потреба у розробці нових інтелектуальних моделей, методів та інформаційних технологій, які б враховували специфіку соціально-економічних даних ТГ, забезпечували високу точність прогнозування, адаптивність до змінних умов та у прийнятті рішень.

Аналіз сучасного стану досліджень та практики обробки соціально-економічних

даних ТГ виявив низку суттєвих протиріч, що обмежують можливості використання цих даних для управління та прийняття рішень. Зокрема, існують такі протиріччя:

- між швидким зростанням обсягів соціально-економічних даних, які є гетерогенними та динамічними, і обмеженими можливостями сучасних методів їх інтеграції, аналізу та прогнозування. Традиційні інструменти не здатні адаптуватися до постійних змін соціально-економічних процесів і не враховують всю різноманітність джерел даних;

- між необхідністю швидкого реагування на сучасні виклики, такі як пандемії, економічна нестабільність чи військова агресія, і відсутністю інтегрованих інформаційних систем, що забезпечують точний аналіз і оперативне прогнозування. Наявні технології часто мають низьку адаптивність до кризових умов, що знижує їхню точність у реальному часі;

- між вимогою до інтеграції різнорідних джерел даних для забезпечення комплексного аналізу (наприклад, статистичних показників, відкритих даних, соціальних опитувань) і технічними бар'єрами їхньої обробки. Такі бар'єри включають проблеми сумісності, нерівномірності даних та високі обчислювальні витрати;

- між потребою в прозорості та зрозумілості результатів аналізу для прийняття обґрунтованих рішень і складністю існуючих методів, які часто незрозумілі для кінцевих користувачів, включаючи управлінців ТГ.

Для розв'язання зазначених протиріч у дисертаційній роботі сформульовано актуальну *науково-прикладну проблему* щодо розроблення методологічних основ інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ. Ця технологія забезпечить адаптивність до змінних умов за рахунок використання методів машинного навчання для моделювання нестационарних і динамічних процесів; інтеграцію різнорідних джерел даних через створення інструментів для збирання, обробки та аналізу структурованих і неструктурованих даних; прозорість і зрозумілість завдяки моделям з пояснюваними результатами, що сприятимуть прийняттю управлінських рішень, а також оперативність і точність прогнозування через впровадження адаптивних прогнозних моделей з ковзним вікном для роботи в режимі реального часу.

Зв'язок роботи з науковими програмами, планами, темами, грантами. Наведені в дисертації дослідження проводились в рамках міжнародних та державних науково-дослідних проєктів, зокрема: Erasmus+: KA2 CBHE «Міждоменні компетенції для забезпечення здорової та безпечної роботи у 21 столітті» (WORK4CE) 619034-EPP-1-2020-1-UA-EPPKA2-CBHE-JP; Creative Europe: AURA – Auralisation of Acoustic Heritage Sites using Augmented and Virtual Reality (nr. 101008547); ERASMUS-EDU-2023-PI-ALL-INNO під назвою "My Farm" (Proposal number: 101140288); держбюджетної наукової роботи молодих учених (номер держреєстрації 0117U003871) «Методи та засоби структурно-статистичної ідентифікації ієрархічних об'єктів за характерними точками їх контурів» (термін виконання: 01.01.2019 р. – 31.12.2020 р.); держбюджетної наукової роботи молодих учених (номер держреєстрації 0117U003871) «TruScanAI: інструмент виявлення фейкової інформації на основі технологій ШІ для боротьби з дезінформацією» (термін

виконання: 01.01.2024 р. – 31.12.2025 р.); науково-дослідної роботи ІОСУ2023 «К» «Інтелектуальні методи, моделі та технології соціально-економічного розвитку ТГ в умовах сьогодення» (протокол № 4 від 09.11.2022 р., термін виконання: 01.01.2023 р. – 31.12.2027 р.); перспективного плану розвитку наукового напрямку «Технічні науки» Західноукраїнського національного університету (номер державної реєстрації 0121U114705, термін виконання: 01.04.2023 р. – 31.12.2023 р.).

Отже, **метою дисертаційної роботи** є забезпечення соціальної стійкості соціально-економічної інфраструктури ТГ шляхом реалізації управлінських рішень на основі інформаційної технології інтелектуального аналізу даних.

Досягнення цієї мети зумовило потребу теоретичних розробок, визначення та послідовного вирішення таких задач:

- 1) провести дослідження сучасного стану та визначити перспективи розвитку аналізу соціально-економічних даних у межах інфраструктури ТГ;
- 2) розробити узагальнений принцип синтезу інформаційної технології для інтелектуального аналізу соціально-економічних даних ТГ;
- 3) розробити методологію аналізу соціально-економічних даних громад, що включає адаптивні методи класифікації, кластеризації та прогнозування;
- 4) розробити новий метод формування навчальної вибірки для нестационарних даних на основі RFM та кластерного аналізу;
- 5) розробити метод ансамблевих адаптивних прогнозних моделей для багатовимірного аналізу;
- 6) розробити метод класифікаційного аналізу кількісних соціально-економічних даних;
- 7) розробити метод класифікаційного аналізу текстових соціально-економічних даних;
- 8) розробити метод кластерного аналізу соціально-економічних даних;
- 9) розробити метод прогнозування соціально-економічних даних;
- 10) розробити метод гібридного аналізу соціально-економічних даних;
- 11) оцінити ефективність запропонованих методів та інформаційної технології.

Об'єкт дослідження: процеси обробки інформації ТГ у динамічному середовищі, які впливають на їх розвиток при стійкому управлінні ресурсами.

Предмет дослідження: інтелектуальні методи, моделі та інформаційні технології аналізу соціально-економічних даних ТГ для підвищення точності прогнозування, оптимізації управління ресурсами та забезпечення стійкого розвитку громад.

Методи дослідження. Для досягнення поставленої мети в дисертаційній роботі використовується комплекс теоретичних і практичних методів, спрямованих на вирішення визначених завдань. Теоретичний аналіз застосовано для вивчення існуючих підходів до обробки та аналізу соціально-економічних даних у контексті управління ТГ. Математичне моделювання використовувалося для розробки нових інтелектуальних моделей і методів, зокрема ансамблевих адаптивних предикторів та методів формування навчальної вибірки. Алгоритми МН були залучені для реалізації інтелектуальних методів прогнозування, класифікації та оптимізації в управлінні соціальними, екологічними та бізнес-процесами ТГ. Експериментальне моделювання

та комп'ютерні симуляції проводилися для оцінки запропонованих методів, з використанням реальних та синтетичних даних. Для верифікації і підтвердження практичної значущості результатів застосовувалися статистичні методи аналізу та порівняння з існуючими підходами.

Наукова новизна результатів дисертаційної роботи. Отримано такі нові наукові результати:

1) вперше запропоновано узагальнені принципи синтезу інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ, особливістю якої є здатність інтегрувати різномірні дані, враховувати динамічність соціально-економічних процесів, що забезпечило підвищення соціальної стійкості інфраструктури громад за рахунок інтеграції гетерогенних даних, адаптивного вибору методів аналізу та формування управлінських рішень;

2) вперше розроблено методологію аналізу соціально-економічних даних ТГ як сукупність методів класифікаційного, кластерного та гібридного аналізу, а також прогнозування соціально-економічних даних, особливістю якої є адаптивність для роботи з нестационарними даними, реальним часом для оперативного прийняття рішень та інтеграцією різномірних джерел даних, що дозволило підвищити точність прогнозування, класифікації та прозорість управлінських рішень;

3) розроблено новий метод формування навчальної вибірки для нестационарних процесів на основі RFM та кластерного аналізу, що, на відміну від існуючих підходів, забезпечує сегментацію та кластеризацію, для підвищення точності опрацювання нових вхідних наборів даних у задачах прогнозування, класифікації та аналізу нестационарних соціально-економічних даних;

4) вперше розроблено метод ансамблевих адаптивних прогнозних моделей для багатовимірного аналізу, особливістю якого є оцінювання на «ковзному вікні» та метамоделі першого і другого рівнів на основі оптимізації функції Лагранжа, що забезпечило точність прогнозів багатовимірних нестационарних процесів з врахуванням їхньої динамічності;

5) удосконалено метод класифікаційного аналізу кількісних соціально-економічних даних, який, на відміну від відомих підходів, забезпечує інтеграцію структурованих, неструктурованих і напівструктурованих даних, що дозволило підвищити точність класифікації кількісних показників для реалізації управлінських рішень;

6) удосконалено метод класифікаційного аналізу текстових соціально-економічних даних, який, на відміну від відомих підходів, забезпечує інтеграцію текстових джерел різної структури та їхній аналіз за допомогою технологій обробки природної мови, що дозволило підвищити ефективність виявлення закономірностей для реалізації управлінських рішень;

7) удосконалено метод кластерного аналізу соціально-економічних даних, який, на відміну від відомих підходів, забезпечує ідентифікацію груп об'єктів з подібними характеристиками, що дозволило підвищити точність розподілу ресурсів для реалізації управлінських рішень;

8) удосконалено метод прогнозування соціально-економічних даних, який, на відміну від відомих підходів, забезпечує високоточне прогнозування динамічних

процесів шляхом інтеграції структурованих, неструктурованих і напівструктурованих даних із застосуванням адаптивних методів інтелектуального аналізу, що дозволило підвищити точність прогнозування для реалізації управлінських рішень;

9) удосконалено метод гібридного аналізу соціально-економічних даних, який, на відміну від відомих підходів, забезпечує врахування кількісних та якісних показників завдяки інтеграції різнорідних даних і багаторівневого підходу до аналізу, що дозволило підвищити гнучкість і адаптивність процесу прийняття рішень для реалізації управлінських рішень.

Практичне значення отриманих результатів. Розробка інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ дозволяє подолати існуючі обмеження традиційних методів обробки даних. Запропонована технологія забезпечує адаптивність до динамічних змін соціально-економічних процесів завдяки використанню методів машинного навчання. Це дозволяє оперативно реагувати на зміни в соціально-економічних показниках та враховувати специфіку різнорідних джерел даних, таких як статистичні показники, відкриті дані та результати соціологічних опитувань. Практичне застосування такої технології сприятиме точному прогнозуванню ризиків, що є критично важливим для управління в умовах сучасних викликів.

Інтеграція різнорідних джерел даних у межах запропонованої технології дозволяє вирішувати завдання комплексного аналізу та моделювання стану ТГ. Це забезпечує можливість виявлення тенденцій, прогнозування змін і оцінювання впливу прийнятих рішень на соціально-економічний стан громад. Створення модулів для обробки як структурованих, так і неструктурованих даних спрощує процес інтеграції інформації з різних джерел, усуваючи технічні бар'єри.

Прозорість і зрозумілість результатів аналізу, забезпечені розробленою інформаційною системою з пояснюваними результатами, дозволяє управлінцям ТГ приймати обґрунтовані рішення. Розроблені інструменти сприяють не лише підвищенню точності при аналізі даних, але й покращенню комунікації між зацікавленими сторонами завдяки доступності інформації для кінцевих користувачів. У підсумку, впровадження технології інтелектуального аналізу даних дозволяє ТГ досягати сталого розвитку, підвищувати соціальну стійкість та адаптуватися до кризових умов.

Реалізація результатів та впровадження. Результати дисертаційної роботи впроваджено в освітні процеси Західноукраїнського національного університету для підготовки здобувачів вищої освіти на першому та другому (освітніх) рівнях за спеціальністю «Комп'ютерні науки» через вдосконалення існуючих навчальних курсів: «Інтелектуальний аналіз даних», «Машинне навчання», «Методи та засоби штучного інтелекту», «Інтелектуальна обробка тексту та природної мови», «Методи та засоби обробки зображень», а також для підготовки здобувачів вищої освіти на третьому (освітньо-науковому) рівні за спеціальністю «Комп'ютерні науки» під час викладання дисциплін «Інтелектуальна обробка даних у складних системах» та «Методи обробки й аналізу експериментальних даних».

Результати дисертаційної роботи впроваджено при виконанні науково-дослідних

робіт та проєктів: Erasmus+: KA2 CBHE «Міждоменні компетенції для забезпечення здорової та безпечної роботи у 21 столітті / «Cross-domain competences for health and safe work in the 21st century» (WORK4CE) 619034-EPP-1-2020-1-UA-EPPKA2-CBHE-JP; Creative Europe: AURA – Auralisation of acoustic heritage sites using Augmented and Virtual Reality (nr. 101008547); ERASMUS-EDU-2023-PI-ALL-INNO, під назвою "My Farm" (Proposal number: 101140288); договір БФ/4-2021 від 01.06.2021 р. «Виконання завдань перспективного плану розвитку наукового напряму «Технічні науки» Західноукраїнського національного університету» (номер державної реєстрації 0121U114705); держбюджетна наукова робота молодих учених (номер держреєстрації 0117U003871 «TruScanAI: інструмент виявлення фейкової інформації на основі технологій ШІ для боротьби з дезінформацією»; науково-дослідна робота ІОСУ2023 «К» «Інтелектуальні методи, моделі та технології соціально-економічного розвитку ТГ в умовах сьогодення.

Переважна більшість цих науково–дослідних робіт і проєктів виконувалась під науковим керівництвом д. т. н. проф. Анатолія Саченка, неоцінімі ідеї, консультації і поради якого зробили суттєвий вклад на дану дисертаційну роботу, і якому авторка висловлює глибоку подяку.

Велику подяку авторка висловлює також проф. Карстену Вольфу (Carsten Wolff) за періодичні консультації і дискусії з основних аспектів дисертаційної роботи і положень наукової новизни, що сприяло покращенню якості дисертації та підготовці ключових статей для періодичних міжнародних видань, індексованих наукометричними базами Web of Science і Scopus.

Результати дисертаційної роботи апробовано та впроваджено у ряді підприємств та установ, таких як: ІТ-компанії «МагнетікВан, м. Тернопіль; Департаменті цифрової трансформації Тернопільської обласної державної адміністрації; Великоберезовицькій ТГ; Громадській організації «Прогресивні люди».

Особистий внесок здобувача. Всі положення, які виносяться на захист, належать автору особисто. Роботи [1–4, 8, 11, 13, 15] виконувалися без співавторів.

У роботах, опублікованих у співавторстві, здобувач особисто виступав в ролі головного дослідника, автора ідеї, аналітика даних, автора методології, експериментатора та програміста. Відповідно до пунктів наукової новизни дисертаційної роботи зроблено такі внески: узагальнений принцип синтезу інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ [1, 17, 19]; методологію аналізу соціально-економічних даних громад з адаптивними методами класифікації, кластеризації та прогнозування [2–5, 17, 19]; метод формування навчальної вибірки для нестационарних даних на основі RFM та кластерного аналізу [4, 27]; метод ансамблю адаптивних прогнозних моделей для багатовимірного аналізу [2, 5, 20, 21]; метод класифікаційного аналізу кількісних соціально-економічних даних [1, 7, 9, 10, 16, 19, 25, 31–34, 38, 39, 41, 45, 49]; метод класифікаційного аналізу текстових соціально-економічних даних з використанням технологій обробки природної мови [1, 12–14, 16, 17, 19, 20, 22, 23, 26, 36, 37, 42–44, 47, 48]; метод кластерного аналізу соціально-економічних даних для ідентифікації груп об'єктів з подібними характеристиками [3, 19, 24]; метод прогнозування соціально-економічних даних з інтеграцією адаптивних методів інтелектуального

аналізу [3, 29, 30, 35, 40, 46]; метод гібридного аналізу соціально-економічних даних з врахуванням кількісних та якісних показників [3, 6, 16, 18, 19].

Особистий внесок інших співавторів у спільних публікаціях: у статті [21] Саченко А. (Sachenko A.) виконував керівництво, концептуалізацію та методологію дослідження; Бодяньський Є. (Bodyanskiy Y.) здійснював математичне редагування, наукове консультування, обговорення результатів дослідження, рецензування та коригування рукопису. У статті [36] Загородня Д. (Zahorodnia D.) виконувала обов'язки автора-кореспондента, адміністрування дослідження; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження; Грам'як Р. (Gramyak R.) займався опрацюванням результатів експериментів; Саченко А. (Sachenko A.) здійснював наукове консультування, керівництво проектом, концептуалізацію дослідження. У статті [46] Комар М. (Komar M.) виконував рецензування та коригування рукопису; Доманський В. (Domanskyi V.) працював над візуалізацією результатів дослідження; Мельник Н. (Melnyk N.) займався опрацюванням результатів експериментів. У статті [29] Комар М. (Komar M.) та Савчишин Р. (Savchishyn R.) виконували рецензування та коригування рукопису; Осолінський О. (Osolinskyi O.) спільно із здобувачем працював над оглядом відомих методів та рішень. У статтях [32–34] Десятнюк О. (Desyatnyuk O.), Крисоватий А. (Krytsovatiy A.), Саченко С. (Sachenko S.), Лукасевич-Крутник І. (Lukasevych-Krutnyk I.) виконували рецензування та коригування рукопису; Бадасян А. (Badasian A.) займався опрацюванням результатів експериментів. У статті [41] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Бодяньський Є. (Bodyanskiy Y.) наукове консультування, рецензування та коригування рукопису; Волф К. (Wolff C.) брав участь у методології дослідження, рецензуванні та коригуванні рукопису. У статті [30] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Бодяньський Є. (Bodyanskiy Y.) здійснював математичне редагування, наукове консультування, рецензування та коригування рукопису; Лендюк Т. (Lendiuk T.) виконав підготовку чернетки рукопису; Подчасова Т. (Podchasova T.) виконувала рецензування та коригування рукопису; Кіт І. (Kit I.) займався візуалізацією результатів дослідження. У статті [22] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Бодяньський Є. (Bodyanskiy Y.) здійснював наукове консультування, рецензування та коригування рукопису. У статті [20] Бодяньський Є. (Bodyanskiy Y.) здійснював математичне редагування, наукове консультування, рецензування та коригування рукопису; Івасечко А. (Ivasechko A.) виконувала обов'язки автора-кореспондента, адміністрування дослідження, візуалізацію результатів дослідження; Кустра Н. (Kustra N.) виступала як рецензент. У статті [28] Загородня Д. (Zahorodnia D.) виконувала обов'язки автора-кореспондента, адміністрування дослідження; Кіт І. (Kit I.) працював над візуалізацією результатів дослідження; Осолінський О. (Osolinskyi O.) спільно із здобувачем працював над оглядом відомих методів та рішень. У статті [19] Комар М. (Komar M.) виконував рецензування та коригування рукопису; Комарницький Р. (Komarnytsky R.) та Мельник Н. (Melnyk N.) займалися опрацюванням результатів експериментів. У статті [18] Комар М. (Komar M.)

виконував рецензування та коригування рукопису; Гаврилюк М. (Havryliuk M.) та Шиманський В. (Shymaskyi V.) виступали як рецензенти; Осолінський О. (Osolinskyi O.) спільно із здобувачем працював над оглядом відомих методів та рішень, виконував коригування рукопису. У статті [37] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. (Lendiuk T.) виконав адміністрування дослідження, підготовку чернетки рукопису; Воложин Ю. (Woloszyn J.) виконував рецензування та коригування рукопису. У статті [26] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Загородня Д. (Zahorodnia D.) займалась адміністрування дослідження; Лендюк Т. (Lendiuk T.) виконував підготовку чернетки рукопису; Комар М. (Komar M.) виконував рецензування та коригування рукопису; Осолінський О. (Osolinskyi O.) спільно із здобувачем працював над оглядом відомих методів та рішень. У статті [39] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лукасевич-Крутник І. (Lukasevych-Krutnyk I.) виступала як рецензент; Бутрин-Бока Н. (Butryn--Boka N.) займалась опрацюванням результатів експериментів. У статті [48] Копаня Л. (Korania Ł.) та Майка Н. (Maika N.) виступали як рецензенти; Саченко С. (Sachenko S.) виконувала рецензування та коригування рукопису; Соія М. (Soia M.) займалась опрацюванням результатів експериментів. У статті [45] Гладій Г. (Hladiy G.) виконував коригування рукопису; Мельничук А. (Melnychuk A.) виконувала рецензування; Юрків Х. (Yurkiv K.) займалась опрацюванням результатів експериментів; Телька М. (Telka M.) виконував підготовку чернетки рукопису. У статті [35] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Кораня Л. та Семанюк В. (Semaniuk V.) виступали як рецензенти; Badasian A. займалась опрацюванням результатів експериментів. У статті [43] Івасечко А. (Ivasechko A.) виконував адміністрування дослідження та візуалізацію результатів дослідження; Соія М. (Soia M.) та Юрків Х. (Yurkiv K.) займалися опрацюванням результатів експериментів. У статті [24] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Волф К. (Wolff C.) брав участь у методології дослідження, рецензуванні та коригуванні рукопису; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Гродський С. (Grodskiy S.), Чижовська З. (Chyzhovska Z.) займалися опрацюванням результатів експериментів. У статті [17] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію дослідження, методологію дослідження; Волф К. (Wolff C.) брав участь у методології дослідження, рецензуванні та коригуванні рукопису; Десятнюк О. (Desyatnyuk O.), Саченко С. (Sachenko S.) виконували рецензування та коригування рукопису; Кіт І. (Kit I.) працював над візуалізацією результатів дослідження. У статті [16] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Волф К. (Wolff C.) брав участь у методології дослідження, рецензуванні та коригуванні рукопису; Загородня Д. (Zahorodnia D.) виконувала адміністрування дослідження, підготовку чернетки рукопису; Кіт І. (Kit I.) працював над візуалізацією результатів дослідження. У статті [25] Максимович В. (Maksymovych V.) виконувала

рецензування та коригування рукопису; Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Кіт І. (Kit I.) працював над візуалізацією результатів дослідження. У статті [23] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Саченко О. (Sachenko O.), Васильків Н. (Vasylykiv N.) виконували рецензування та коригування рукопису. У статті [31] Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Саченко С. (Sachenko S.) та Яцків В. (Yatskiv V.) виконували рецензування та коригування рукопису; Брич В. (Brych V.) займався опрацюванням результатів експериментів. У статті [27] Загородня Д. (Zahorodnia D.) виконувала підготовку чернетки рукопису; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження; Саченко С. (Sachenko S.) виконувала рецензування та коригування рукопису; Брич В. (Brych V.) займався опрацюванням результатів експериментів; Биковий П. (Bykovyy P.) працював над візуалізацією результатів дослідження. У статті [42] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Ткачик Ф. (Tkachyk F.), Парій В. (Pariy V.) спільно з автором працювали над оглядом відомих методів та рішень; Брюханський Р. (Brukhanskyi R.), Лендюк Д. (Lendyuk D.) займалися опрацюванням результатів експериментів. У статті [44] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Добровольський М. (Dobrowolski M.) виконував рецензування та коригування рукопису; Богута Г. (Boguta G.), Бицюра Л. (Bytsyura L.) займалися опрацюванням результатів експериментів. У статті [47] Саченко А. (Sachenko A.) виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. (Lendiuk T.) виконував адміністрування дослідження, підготовку чернетки рукопису; Коваль В. (Koval V.), Гладій Г. (Hladiy G.) виконували рецензування та коригування рукопису; Галіас Ю. (Halias Y.) займався опрацюванням результатів експериментів. У статті [7] Комар М. виконував рецензування та коригування рукопису, методологію дослідження; Кіт І. працював над візуалізацією результатів дослідження; Мадараш Р., Юрків Х. займалися опрацюванням результатів експериментів. У статті [14] Кіт І. працював над візуалізацією результатів дослідження. У статті [5] Юрків Х. займалася опрацюванням результатів експериментів. У статті [10] Саченко А. виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. виконував адміністрування дослідження, підготовку чернетки рукопису; Комар М. виконував рецензування та коригування рукопису. У статті [9] Саченко А. виконував керівництво проектом, концептуалізацію та методологію дослідження; Лендюк Т. виконував адміністрування дослідження, підготовку чернетки рукопису; Комар М. виконував рецензування та коригування рукопису. У статті [6] Комар М. виконував рецензування та коригування рукопису; Юрків Х. займалася опрацюванням результатів експериментів; Лукянчук В. спільно з автором працював над оглядом

відомих методів та рішень.

Апробація результатів дисертації. Основні результати дисертаційного дослідження багаторазово доповідалися та обговорювалися на 19 міжнародних і національних конференціях, а саме: 10th, 11th, 12th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (18–21 вересня 2019 р., Мец, Франція; 22–25 вересня 2021 р., Краків, Польща; 7–9 вересня 2023 р., Дортмунд, Німеччина); 4th, 5th IEEE International Conference on Advanced Information and Communication Technologies (AICT) (20–24 вересня 2021 р., Львів, Україна; листопад 2023 р., Київ, Україна); 5th, 7th International Workshop on Computer Modeling and Intelligent Systems (CMIS) (12 травня 2022 р., Запоріжжя, Україна; 3 травня 2024 р., Запоріжжя, Україна); International Workshop on Intelligent Information Technologies & Systems of Information Security (IT&SIS) (23–25 вересня 2021 р., Хмельницький, Україна; 29–30 вересня 2022 р., Хмельницький, Україна); International Workshop on Modern Machine Learning Technologies (MoMLeT) (20–22 вересня 2021 р., Львів, Україна; 31 травня – 1 червня 2024 р., Львів, Україна); International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI) (20–22 листопада 2023 р., Ватерлоо, Канада); 16th IEEE International Conference on Computer Sciences and Information Technologies (CSIT) (22–25 вересня 2021 р., Львів, Україна); IEEE European Technology and Engineering Management Summit (E-TEMS) (9–11 березня 2023 р., Каунас, Литва); First International Workshop of Young Scientists on Artificial Intelligence for Sustainable Development (YS-AISD) (10–11 травня 2024 р., Тернопіль, Україна); International Conference on Computational Linguistics and Intelligent Systems (COLINS) (22–23 квітня 2021 р.; 18–19 квітня 2024 р.); III International Scientific Symposium “Intelligent Solutions” (INTSOL-2023) (26–28 квітня 2023 р.); IEEE International Conference on Problems of Infocommunications, Science and Technology (PIC S&T) (5–7 жовтня 2021 р.); 16th International Conference on ICT in Education, Research and Industrial Applications (ICTERI 2020) (6–8 жовтня 2020 р., Харків, Україна); Advances in Computer Science for Engineering and Education IV, частина Міжнародної конференції з комп'ютерних наук, інженерії та освітніх застосувань (ICCSEEA) (15–17 січня 2021 р.); International Workshop on IT Project Management (ITPM) (14–17 вересня 2021 р.).

Публікації. За темою дисертації опубліковано 49 наукових праць, зокрема: 15 статей – у наукових виданнях, включених до Переліку наукових фахових видань України [1–15]; 7 статей – у періодичних виданнях, включених до категорії «А» Переліку наукових фахових видань України, або у закордонних виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus [16–48]. Відповідно до класифікації SCImago Journal and Country Rank, 1 статтю віднесено до квартилю Q1 [18], 4 статті – до квартилю Q2 [16–20], 1 стаття – до квартилю Q4 [21]; 4 монографії (розділи у колективних монографіях) [23–26]; 22 статті – у матеріалах міжнародних конференцій, які індексуються у наукометричних базах Scopus та Web of Science [27–48]; 1 публікація, яка додатково відображає наукові результати дисертації [49].

Структура та обсяг дисертації. Дисертаційна робота складається з анотацій, вступу, 7 розділів, висновків, 5 додатків на 45 сторінках та списку використаних джерел з 475 найменувань на 52 сторінках. Загальний обсяг дисертації становить 361 сторінок, з них 299 сторінок основного тексту, 76 рисунків та 15 таблиць.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертації, визначено об'єкт та предмет, сформульовано мету та задачі дослідження, представлено наукову новизну та практичну цінність одержаних результатів. Зазначено зв'язок роботи з науковими програмами та темами за місцем виконання роботи. Надано інформацію щодо кількості публікацій та апробації результатів дисертації.

У **першому розділі** здійснено аналіз сучасного стану й перспектив розвитку аналізу соціально-економічних даних в інфраструктурі ТГ. Виокремлено ключові аспекти, що впливають на ефективність прийняття управлінських рішень, а саме: гетерогенність та динамічність даних, необхідність моделювання соціально-економічної інфраструктури, інтеграція різних джерел інформації та використання інтелектуальних методів аналізу.

Встановлено, що традиційні підходи до аналізу соціально-економічної інформації не забезпечують оперативного реагування на дестабілізуючі фактори, такі як пандемія, економічна нестабільність та військова агресія. Аналіз літературних джерел засвідчив відсутність цілісних рішень, здатних одночасно враховувати кількісні та якісні показники, підтримувати прийняття адаптивних управлінських рішень, а також інтегрувати результати з різних методів (кластеризації, класифікації, прогнозування).

Виявлено, що однією з перспективних стратегій є впровадження інтелектуальних інформаційних технологій, заснованих на машинному навчанні та сучасних підходах до обробки неструктурованої інформації. Зокрема, застосування бустингових алгоритмів і ансамблевих методів, а також використання кластерного та класифікаційного аналізу й адаптивного прогнозування, може суттєво підвищити точність, надійність і швидкість ухвалення управлінських рішень у ТГ.

Отримані результати свідчать про необхідність розробки комплексної інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ, яка дозволить інтегрувати дані з різних джерел, обробляти їх з урахуванням нестаціонарності та гетерогенності, а також забезпечити стає функціонування інфраструктури громади. У такій технології основними компонентами виступатимуть моделі, бази знань, правила та методи обробки даних, що дозволять формувати релевантні управлінські альтернативи.

Здійснено порівняльний аналіз основних підходів і методів інтелектуального аналізу, запропоновано застосування бустингових алгоритмів для підвищення точності прогнозів, а також обґрунтовано необхідність використання різних метрик оцінки якості (наприклад, точності, F1-міри, AUC-ROC) для достовірної перевірки отриманих результатів. Особливу увагу приділено адаптації методів до умов незбалансованих даних, вагових коефіцієнтів та інших спеціалізованих підходів.

Таким чином, результати аналізу підтверджують актуальність та важливість розробки та застосування інтелектуальної інформаційної технології аналізу соціально-економічних даних у ТГ, що сприятиме підвищенню соціальної стійкості, оптимізації використання ресурсів, покращенню якості життя мешканців та ефективному реагуванню на сучасні виклики, пов'язані з економічною нестабільністю, екологічними ризиками та військовою агресією.

У **другому розділі** розроблено узагальнену інформаційну технологію інтелектуального аналізу соціально-економічних даних ТГ, яка спрямована на

підвищення їхньої стійкості в кризових умовах, таких як пандемія чи війна. В основу технології покладено ідею інтеграції гетерогенних даних з різних джерел, визначення ключових об'єктів інфраструктури та використання адаптивних методів інтелектуального аналізу для підтримки стратегічних управлінських рішень. Окрім цього, у розділі представлено методологію адаптивного аналізу, яка забезпечує ефективну роботу з нестационарними даними та підвищення точності прогнозування, а також запропоновано методи формування навчальних вибірок і ансамблю адаптивних прогнозних моделей для багатовимірних нестационарних послідовностей в реальному часі.

Узагальнена інформаційна технологія інтелектуального аналізу соціально-економічних даних ТГ [1] спрямована на підвищення їхньої стійкості в кризових умовах, таких як пандемія чи війна. Суть підходу полягає у визначенні ключових об'єктів інфраструктури (I), інтеграції гетерогенних даних (D) з різних джерел та застосуванні адаптивних методів інтелектуального аналізу (A) для підтримки прийняття стратегічних рішень. На рис. 1 представлено узагальнену структуру цієї технології, де між об'єктами інфраструктури, типами даних та методами аналізу існує взаємозв'язок, що дозволяє підвищити стійкість ТГ у кризових ситуаціях.

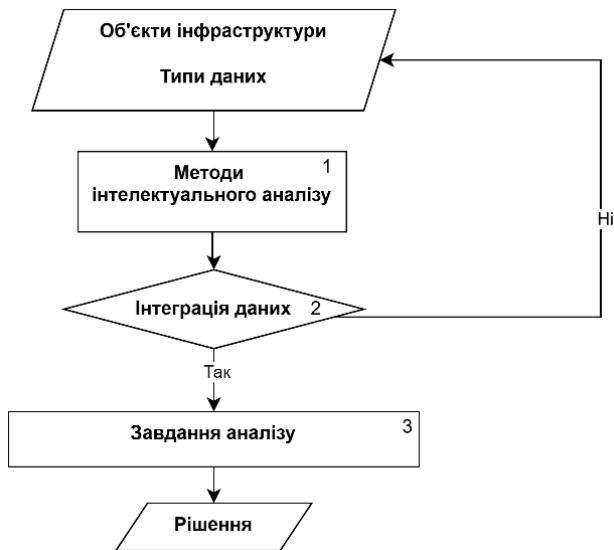


Рис. 1. Узагальнена структура інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ

В основі формалізації підходу лежить поділ множин I, D та A . Нехай $I = \{I_1, I_2, \dots, I_n\}$ – множина об'єктів інфраструктури, $D = \{D_1, D_2, D_3\}$ – множина типів даних (структуровані, неструктуровані та напівструктуровані), $A = \{A_1, A_2, \dots, A_m\}$ – множина методів інтелектуального аналізу (машинне навчання, обробка природної мови, комп'ютерне бачення тощо). Функція застосовності $f: D \times A \rightarrow \{0,1\}$ визначає, чи метод A_k можна застосувати до типу даних D_i .

Якщо $f(D_i, A_k) = 1$, тоді метод аналізу сумісний з цим типом даних.

Для інтеграції даних використовується оператор об'єднання:

$$D_{\text{інтегровані}} = D_1 \oplus D_2 \oplus \dots \oplus D_n, \quad (1)$$

де $D_{\text{інтегровані}}$ є підмножиною об'єднання всіх типів даних, залучених із різних джерел.

Кожне завдання аналізу формалізується як трійка

$$Z_l = (I_j, D_i, A_k), \quad (2)$$

що відображає вибір конкретного об'єкта інфраструктури, типу даних та методу аналізу з урахуванням сумісності, визначеної функцією f .

Завдяки такій формальній моделі можливо гнучко підбирати методи аналізу для об'єктів інфраструктури ТГ, інтегрувати різні типи даних та забезпечувати розробку рішень $R_p = g(Z_l)$, спрямованих на підвищення стійкості громади. Таким чином, запропонована інформаційна технологія дозволяє вчасно виявляти актуальні проблеми, прогнозувати розвиток ситуацій та приймати стратегічні рішення, що сприяють ефективному управлінню соціально-економічною інфраструктурою ТГ.

Методологія інтелектуального аналізу соціально-економічних даних територіальних громад [2]. Запропонована методологія інтелектуального аналізу соціально-економічних даних ТГ спрямована на підвищення точності прогнозування, адаптивності управлінських рішень та оптимізації процесів у динамічному середовищі. Вона поєднує принципи адаптивності, інтеграції та прозорості із застосуванням сучасних підходів машинного навчання, методів аналізу в режимі реального часу та ансамблевих адаптивних предикторів. Завдяки цьому забезпечується можливість ефективно інтегрувати різномірні джерела даних, швидко реагувати на зміни та підтримувати стійкий розвиток ТГ в умовах кризи.

Основу методології складають кілька послідовних етапів (рис. 2): визначення задачі аналізу, збір релевантних даних, формування навчальних вибірок, застосування інтелектуальних методів, використання ансамблевих адаптивних предикторів та аналіз результатів. Кожен етап відповідає певній категорії управлінських завдань: від підтримки соціальної стійкості та розвитку (оцінка якості життя, формування споживчого кошика, боротьба з дезінформацією) до управління підприємницькою діяльністю (оцінка інвестиційних ризиків, прогнозування попиту, виявлення фіктивних підприємств).

Окрему увагу приділено принципам, сформульованим на основі протиріч, притаманних сучасному аналізу даних. Принцип адаптивності дозволяє зберігати точність за умов швидкої зміни даних; принцип аналізу в реальному часі дає змогу оперативно реагувати на виклики; принцип інтеграції об'єднує різномірні джерела інформації; принцип точності прогнозування підсилює результати за допомогою передових підходів машинного навчання; а принцип прозорості забезпечує зрозуміле представлення висновків для всіх зацікавлених сторін.

Реалізація методології підтверджена опублікованими науковими працями, що засвідчують її ефективність у підвищенні точності прогнозів, адаптивності аналізу та інтерпретації результатів. Інтеграція передових алгоритмів, адаптивних підходів і гнучких інструментів аналітики дозволяє оптимізувати управління ресурсами ТГ та приймати обґрунтовані рішення навіть у динамічних і кризових умовах.

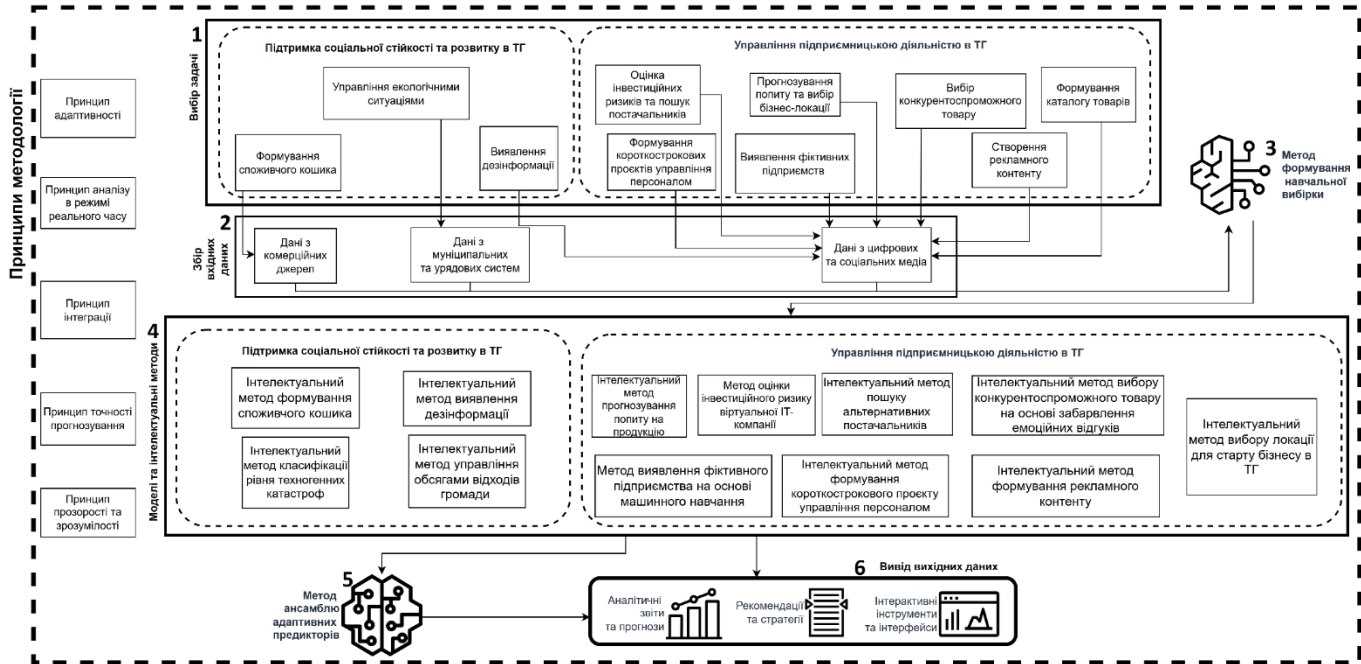


Рис. 2. Структура методології аналізу соціально-економічних даних ТГ

Метод формування навчальної вибірки для нестационарних даних на основі RFM та кластерного аналізу [2, 4]. Представлений метод формування навчальної вибірки для нестационарних даних на основі RFM та кластерного аналізу дозволяє автоматично адаптувати вибірку під зміни вхідної інформації та відбирати найбільш репрезентативні об'єкти. Ключовою відмінністю цього підходу є використання показників Recency (R), Frequency (F), Monetary (M) для формування первинних сегментів, а далі – застосування кластерного аналізу (зокрема, методу k -середніх) для визначення та уточнення груп. Цей процес більш наочно зображено на рис. 3, де показана покрокова структура формування навчальної вибірки.



Рис. 3. Структурно-логічна схема методу формування навчальної вибірки

Першим етапом є розрахунок RFM показників для кожного об'єкта:

$$RFM_i = (R_i, F_i, M_i), \quad (3)$$

де R_i – давність останньої транзакції; F_i – частота транзакцій; M_i – загальна сума транзакцій для i -го об'єкта.

Після цього здійснюється розподіл значень R, F, M за кuartилями, присвоюються ранги, а потім об'єкти групуються за сумарним RFM-індексом.

Наступним кроком є пониження розмірності даних з використанням методу головних компонент (РСА), що сприяє скороченню вимірності ознак та покращенню якості подальшої кластеризації. Застосування методів «лікоть» та силуетного аналізу для вибору оптимальної кількості кластерів допомагає підвищити якість та стабільність сегментації.

Основною формулою для обчислення коефіцієнта силуету, який дозволяє оцінити якість кластеризації, є:

$$s_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}, \quad (4)$$

де a_i – середня відстань від точки i до інших точок свого кластера; b_i – мінімальна середня відстань від точки i до точок найближчого іншого кластера.

Після визначення оптимальної кількості кластерів за допомогою коефіцієнта силуету та ініціалізації кластерних центрів методом K–Means, відбувається формування чітких, стійких груп. Ці групи та їх розподіл характеризуються результатами кластеризації об'єктів на основі їхніх RFM-оцінок. Застосування запропонованого підходу дозволяє виділити сегменти, такі як «Найкращі організатори», «Лояльні організатори», «Великі споживачі» тощо, що сприяє більш чіткій структуризації даних для подальшого аналізу.

Отримані кластери використані для формування навчальної вибірки, де до початкових ознак додаються рангові оцінки RFM, координати в просторі головних компонент та мітки кластерів. Це перетворює дані на більш структуровану та інформативну вибірку для навчання моделей машинного навчання. Результати показують, що застосування зазначеного методу дозволяє підвищити точність класифікації об'єктів (наприклад, організаторів тендерних проєктів) за допомогою алгоритмів Logistic Regression та Naive Bayes до рівня 99–100 %.

Перевага описаного підходу підтверджена результатами, опублікованими в авторитетних наукових джерелах. Метод дає можливість адаптивно реагувати на зміну характеристик вхідних нестаціонарних даних, виділяти та враховувати різні типи поведінкових шаблонів, знижувати вплив шуму та неінформативних ознак. Завдяки цьому створюється більш якісна навчальна вибірка, що підвищує ефективність побудови моделей прогнозування та класифікації у різних галузях, зокрема при аналізі великих соціально-економічних масивів даних.

Метод ансамблю адаптивних прогнозних моделей для багатовимірних нестаціонарних послідовностей в реальному часі [2, 21]. Метод ансамблю адаптивних прогнозних моделей розроблено для розв'язання задачі прогнозування багатовимірних нестаціонарних послідовностей у реальному часі, що є особливо актуальним в умовах динамічних змін сигналів і високої невизначеності. Традиційні підходи, такі як регресія чи нейронні мережі, часто демонструють недостатню точність через переоснащення або недооснащення, а також не враховують змінність характеристик сигналу. Використання ансамблів прогнозуючих моделей з адаптивним налаштуванням параметрів дозволяє суттєво підвищити точність прогнозів.

Метод базується на використанні ансамблю багатовимірних предикторів (MP_1, \dots, MP_h), кожен з яких генерує прогнози $\hat{x}_j(\tau)$, для одного і того самого сигналу $x(\tau)$. Об'єднаний прогноз формується метамодделлю другого рівня за формулою:

$$x^*(\tau) = \sum_{j=1}^h c_j \hat{x}_j(\tau), \quad (5)$$

де c_j – ваги, визначені методом оптимізації, які забезпечують баланс між прогнозами різних моделей.

Для цього використовується класичний метод Лагранжа, що мінімізує квадратичну похибку прогнозу за умови, що:

$$\sum_{j=1}^h c_j = 1. \quad (6)$$

Параметри метамоделі c_j обчислюються шляхом розв'язання рівнянь Куна–Таккера, що дозволяють знайти оптимальні ваги для кожного предиктора. Основна функція оптимізації виглядає так:

$$L(c, \lambda) = Sp(V^T(T)V(T)) + \lambda(c^T E_h - 1), \quad (7)$$

де $V(T)$ – матриця похибок прогнозів; E_h – вектор одиниць; λ – множник Лагранжа.

Остаточне рішення для ваг:

$$c = c^* + D(T) \frac{1 - E_h^T c^*}{E_h^T D(T) E_h} E_h, \quad (8)$$

де $D(T)$ – матриця, обернена до коваріаційної матриці прогнозів.

Для обробки нестационарних сигналів метод використовує процедури оцінювання на основі «ковзного вікна», де розмір вікна s визначає кількість останніх спостережень, які використовуються для оцінки. Формули для адаптації параметрів метамоделі включають:

$$D(T+1) = D(T) - D(T) \hat{x}^T(T+1) \times (E_{nn} + \hat{x}(T+1)D(T)\hat{x}^T(T+1))^{-1} \times \hat{x}(T+1)D(T). \quad (9)$$

Такий підхід дозволяє постійно оновлювати ваги метамоделі залежно від останніх змін у даних.

Ефективність методу продемонстровано на ансамблі моделей, до якого входили логістична регресія, дерева рішень, метод опорних векторів, випадковий ліс та градієнтний бустинг. Метамоделю другого рівня побудована з використанням алгоритму адаптивного підсилення AdaBoostClassifier. Згідно з експериментальними даними, точність прогнозування досягла 98 %, F1-оцінка становила 0.98 для одного з класів, що демонструє перевагу методу над окремими моделями.

Метамоделю забезпечує комбінування переваг кожного члена ансамблю, враховуючи їхню точність та узгодженість, що дозволяє покращити загальну продуктивність. Додатково, адаптивна структура гарантує стійкість до змін характеристик сигналів і ефективність навіть у складних динамічних умовах. Запропонований метод (рис. 4) дозволяє також уникати переоснащення моделі, зберігаючи її здатність до узагальнення.

Таким чином, представлені підходи, зокрема узагальнена інформаційна технологія аналізу соціально-економічних даних, метод формування навчальної вибірки для нестационарних даних та метод ансамблю адаптивних прогнозних моделей, демонструють високу ефективність у розв'язанні складних задач прогнозування та управління. Вони забезпечують адаптивність до динамічних змін даних, інтеграцію гетерогенних джерел і використання сучасних алгоритмів аналізу, таких як машинне навчання, кластеризація та метамоделі. Завдяки цьому стає можливим підвищення стійкості ТГ, покращення якості прогнозів і прийняття стратегічних рішень навіть у кризових умовах.

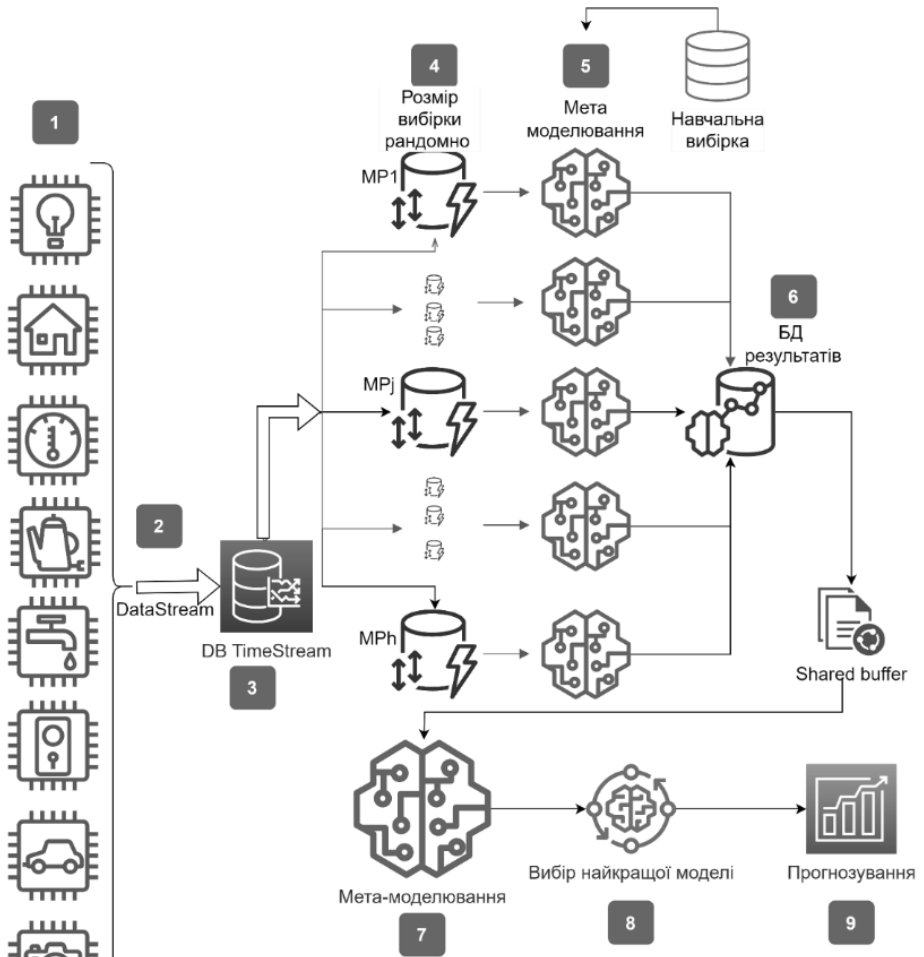


Рис. 4. Структурно-логічна схема методу ансамблю адаптивних прогнозних моделей для багатовимірних нестационарних послідовностей в реальному часі

У **третьому розділі** розроблено метод класифікаційного аналізу кількісних соціально-економічних даних ТГ [1] (рис. 5), який ґрунтується на узагальненому принципі синтезу інформаційних технологій інтелектуального аналізу даних (див. рис. 1). Важливим кроком цього методу є інтеграція різномірних типів даних, що виражається рівнянням (1). Наступним етапом є формування завдань аналізу, яке представлено трійкою:

$$Z_l = (I_j, D_i, A_k), \quad (10)$$

де I_j – об'єкти інфраструктури (бізнес-інфраструктура I_1 та економічна інфраструктура I_6); D_i – типи даних (D_1, D_2, D_3); A_k – методи аналізу (Logistic Regression, SVM, Random Forest, kNN, Gaussian Naive Bayes).

Ця структура дозволяє точно поєднувати об'єкти інфраструктури з відповідними типами даних і методами аналізу, забезпечуючи гнучкість та адаптивність методу до різних завдань, таких як оцінка інвестиційних ризиків чи виявлення фіктивних підприємств.



Рис. 5. Кроки методу класифікаційного аналізу кількісних соціально-економічних даних ТГ

Фінальним кроком є прийняття рішень на основі аналізу, яке описується функцією: $R_p = g(Z_1)$, де g – функція вибору оптимального рішення, наприклад, визначення найкращого інвестиційного проекту або класифікація підприємства як фіктивного. Ця функція використовує результати класифікаційного аналізу для формулювання конкретних рекомендацій, спрямованих на підвищення соціально-економічної стійкості ТГ. Завдяки інтеграції різних типів даних та застосуванню передових методів машинного навчання, розроблений метод класифікаційного аналізу є ефективним інструментом для управління ризиками і підтримки прийняття рішень у рамках ТГ.

На основі запропонованого методу класифікаційного аналізу кількісних соціально-економічних даних ТГ у розділі детально представлено підходи до оцінки інвестиційного ризику віртуальної ІТ-компанії на основі методів машинного навчання [9, 25] та метод виявлення фіктивного підприємства навчання [31–34, 49]. Розроблений метод класифікаційного аналізу став основою для низки наукових досліджень, зокрема, в [16, 29, 39, 41, 45], де підтверджено його ефективність і адаптивність у вирішенні завдань управління ризиками та покращення соціально-економічної стійкості ТГ.

Реалізація оцінки інвестиційного ризику для віртуальної ІТ-компанії заснована на алгоритмі машинного навчання, зокрема Support Vector Classifier (SVC), Random Forest Classifier та K-Nearest Neighbors (kNN), що забезпечують високий рівень точності прогнозування. Згідно з таблицею 1, методи класифікації демонструють такі результати: SVC досяг точності прогнозу 0,982 з середнім квадратичним відхиленням 0,014, Random Forest Classifier – 0,984, з відхиленням 0,021; kNN – 0,981, з відхиленням 0,024. Ці показники підтверджують ефективність застосованих алгоритмів у визначенні оптимальних інвестиційних проектів, де обидва методи (SVC та Random

Forest) вибрали проєкт з кодом P8 як найбільш привабливий з точністю 0,98.

Кількісні результати (табл. 1) демонструють високу ефективність запропонованого методу, забезпечуючи точність класифікації на рівні 98 % для всіх трьох моделей. Наприклад, Random Forest Classifier показав середню абсолютну похибку MAE = 0,05 та точність Score = 0,98, що підтверджує його здатність точно ідентифікувати проєкти з низьким рівнем ризику. Крім того, KNeighborsClassifier також продемонстрував високу точність 0,98, хоча з нижчою точністю для проєктів, які варто інвестувати (0,67). Загалом, результати оцінки інвестиційних ризиків свідчать про здатність системи швидко та точно аналізувати великі обсяги даних, що суттєво скорочує час ухвалення рішень від 1–5 годин до 1–3 днів без використання системи. Таким чином, запропонований метод не лише мінімізує інвестиційні ризики, але й підвищує стабільність розвитку віртуальних ІТ-компаній у межах ТГ.

Таблиця 1

Результати перехресної оцінки

Метод	Оцінка прогнозу	Середнє квадратичне відхилення
1. SupportVectorClassifier	0,982	0,014
2. StochasticGradientDecentC	0,963	0,024
3. RandomForestClassifier	0,984	0,021
4. DecisionTreeClassifier	0,951	0,042
5. GaussianNB	0,962	0,023
6. KNeighborsClassifier	0,981	0,024
7. AdaBoostClassifier	0,953	0,044
8. LogisticRegression	0,984	0,027

Для виявлення фіктивних підприємств було застосовано метод класифікації за допомогою машинного навчання на основі аналізу даних 100 компаній з України, 20 з яких були визначені як фіктивні. Використовуючи 8 різних алгоритмів класифікації, серед яких Logistic Regression, Support Vector Classifier та Gaussian Naive Bayes, досягнуто високі показники точності прогнозу до 0,98 з мінімальним стандартним квадратичним відхиленням 0,02 (табл. 2). Найкращі результати були отримані саме цими трьома методами, що свідчить про їх ефективність у задачі виявлення фіктивних підприємств. Таблиця 2 демонструє порівняльні результати всіх використаних алгоритмів, підкреслюючи перевагу Logistic Regression, Support Vector Classifier та Gaussian Naive Bayes.

Таблиця 2

Результати перехресної оцінки

№	Метод	Оцінка прогнозу	Середнє квадратичне відхилення
1	SupportVectorClassifier	0,983333	0,023570
2	StochasticGradientDecentC	0,966667	0,023570
3	RandomForestClassifier	0,966667	0,023570
4	DecisionTreeClassifier	0,950000	0,040825
5	GaussianNB	0,983333	0,023570
6	KNeighborsClassifier	0,883333	0,062361
7	AdaBoostClassifier	0,950000	0,040825
8	LogisticRegression	0,983333	0,023570

Впровадження розробленої інформаційної системи значно скоротило час, необхідний для виявлення фіктивних підприємств з 3–20 днів до 2–6 годин, забезпечуючи ефективність на рівні 99,58 %. Крім того, модель логістичної регресії показала високий показник $AUC \approx 0,99$, що підтверджує її високу здатність до розрізнення між фіктивними і реальними підприємствами. Це дозволяє користувачам швидко та точно приймати рішення щодо співпраці з потенційними партнерами, мінімізуючи ризики шахрайства. Загалом, дослідження демонструє, що застосування методів машинного навчання, зокрема Logistic Regression та Support Vector Machine, є надзвичайно ефективними інструментами для інтелектуального виявлення фіктивних підприємств у сучасному електронному бізнес-середовищі.

Отже, розроблено та впроваджено метод класифікаційного аналізу соціально-економічних даних, який використовує різні алгоритми машинного навчання для оцінки інвестиційних ризиків віртуальних ІТ-компаній та виявлення фіктивних підприємств з високою точністю до 99,7 %. Розроблена інформаційна система дозволяє суттєво скоротити час прийняття рішень з кількох днів до кількох годин, забезпечуючи оперативність та точність в аналізі даних. Отримані результати демонструють практичну цінність методів для підвищення ефективності управління ризиками, покращення інвестиційного клімату та зміцнення економічної безпеки ІТ.

У **четвертому розділі** розроблено метод класифікаційного аналізу текстових соціально-економічних даних ІТ [1] (рис. 6), який ґрунтується на узагальненому принципі синтезу інформаційних технологій інтелектуального аналізу даних (див. рис. 1). Основною є інтеграція різнорідних типів даних, виражена рівнянням (1), що відрізняється від попереднього методу тим, що зосереджується на об'єднанні структурованих, неструктурованих та напівструктурованих текстових даних для забезпечення більш повного і гнучкого аналізу. На відміну від кількісного методу, який використовував інфраструктурні об'єкти I_1 та I_6 , новий метод включає додаткові об'єкти, такі як інформаційна (I_1), соціальна (I_2) та освітня інфраструктура (I_3), що дозволяє більш детально охоплювати специфіку текстових даних.



Рис. 6. Кроки методу класифікаційного аналізу текстових соціально-економічних даних ІТ

Крім того, зазначений метод використовує оновлену трійку завдань $Z_l = (I_j, D_i, A_k)$, де методи аналізу A_k включають не лише традиційні алгоритми машинного навчання, але й глибокі нейронні мережі, такі як LSTM і трансформери. Це

дозволяє більш точно класифікувати тексти за рівнем достовірності та емоційного забарвлення, а також визначати важливість факторів для прийняття рішень. Функція прийняття рішень $R_p = g(Z_l)$ адаптована до специфіки текстових даних, забезпечуючи оптимальні результати для задач, таких як автоматизація HR-проектів та ранжування конкурентоспроможних товарів.

На основі методу класифікаційного аналізу текстових соціально-економічних даних ТГ у розділі детально представлені інтелектуальні підходи до виявлення дезінформації [7, 17, 20, 44,] та формування короткострокового HR-проекту [8, 23]. Цей метод, який інтегрує сучасні технології інтелектуального аналізу тексту, став основою для кількох наукових публікацій, зокрема в [3, 9, 12–14, 16–19, 22, 24, 26, 36, 37, 42, 47], що підтверджують його ефективність у вирішенні соціально-економічних завдань ТГ.

Запропонований метод виявлення дезінформації, базуючись на унікальному наборі даних з 60 тисяч заголовків новин періоду російсько-української війни, демонструє високу ефективність в розв'язанні задач класифікації. Використання ансамблевого підходу на основі стекінгу, зокрема метамоделі XGBoost, забезпечило загальну точність класифікації на рівні 93 %. Точність для класу «False» склала 0,95, що свідчить про високу вірогідність правильної класифікації фейкових новин. Однак показник відновлення для цього класу дорівнює 0,72, вказуючи на пропуск близько 28 % фейкових новин. Для класу «True» модель досягла вражаючого результату – 99 % правильного визначення правдивих новин. F1-оцінка, що поєднує точність і відновлення, дорівнює 0,82 для класу «False» та 0,95 – для класу «True», що підтверджує збалансованість методу.

На рис. 7 представлено матрицю помилок, яка візуалізує результати класифікації. Модель коректно ідентифікувала 1793 фейкових новини (True Negative) та 8138 правдивих (True Positive), але припустила 705 хибно негативних (False Negative) та 99 хибно позитивних (False Positive) класифікацій. Це вказує на необхідність подальшого вдосконалення моделі для зменшення пропусків серед фейкових новин. Система аналізу текстових даних, побудована на цьому методі, дозволяє зменшити час перевірки інформації з кількох днів до декількох годин, досягаючи середньої ефективності на рівні 99,17 %. Такий результат робить її потужним інструментом у боротьбі з дезінформацією.

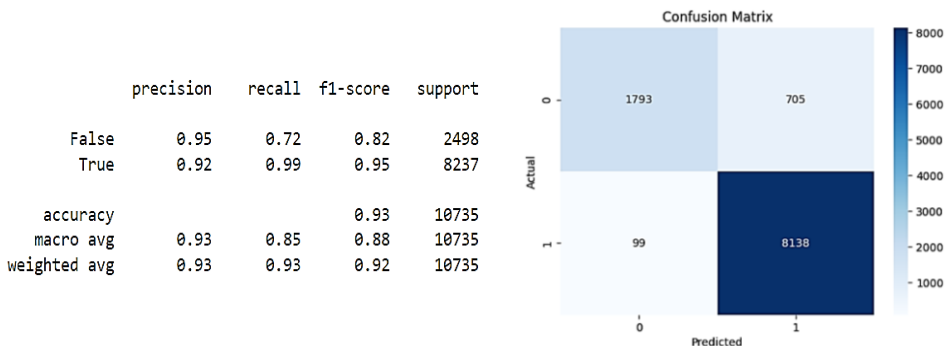


Рис. 7. Результати оцінювання методу виявлення дезінформації

Розроблений метод формування HR-проектів з використанням інтелектуального чат-бота на платформі Telegram продемонстрував значну ефективність у скороченні часу на пошук і відбір кандидатів. Система реалізована через поєднання бібліотек Python, методів машинного навчання (RandomForestClassifier), та інтеграцію з платформою DialogFlow для автоматизації співбесід. Чат-бот забезпечив автоматичне опитування претендентів за трьома категоріями: загальні питання, м'які та тверді навички, з подальшою обробкою відповідей у структуровану базу даних. Використання дерева рішень для класифікації кандидатів дозволило досягти точності на рівні 94 %, а F1-оцінка для визначення відповідності претендента конкретній посаді склала 0,87 для класу «розробник». За допомогою запропонованої системи час підбору кандидатів скоротився з 5–30 днів до 1–5 годин.

На рис. 8 наведено структуру дерева рішень для вибору найкращого кандидата на позицію розробника. Результати аналізу підтверджують, що середній час інтерв'ювання з використанням системи варіюється від 1 до 120 годин залежно від складності проекту та вимог, що значно оптимізує процеси HR-менеджменту. Порівняно з традиційними методами, ефективність системи складає 80 %, що підтверджує її доцільність для короткострокових HR-проектів. Отримані дані не лише автоматизують процес прийняття рішень, але й інтегрують нову інформацію для подальшого навчання моделі, що підвищує її точність та адаптивність у майбутньому.

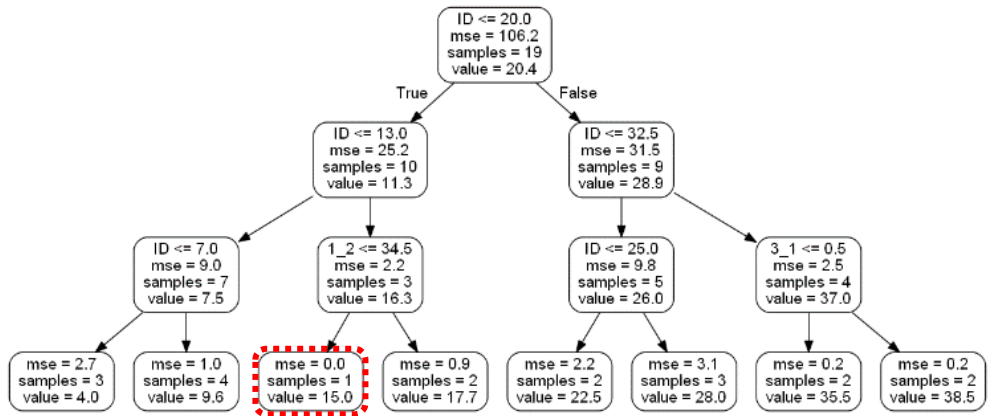


Рис. 8. Дерево рішень для вибору найкращого кандидата на посаду розробника

Отже, представлено метод класифікаційного аналізу текстових соціально-економічних даних ТГ, який інтегрує різноманітні джерела інформації та використовує сучасні алгоритми машинного навчання. Проведені експериментальні дослідження продемонстрували високу ефективність розроблених підходів: метод виявлення дезінформації досяг точності 93 %, а система формування HR-проектів скоротила час виконання завдань з 5–30 днів до 1–120 годин. Запропоновані рішення забезпечують автоматизацію процесів їх прийняття та підвищують якість управління, сприяючи стійкому розвитку соціально-економічних процесів у ТГ.

У п'ятому розділі запропонований метод кластерного аналізу соціально-економічних даних ТГ [3] (рис. 9), який ґрунтується на узагальненому принципі синтезу інформаційних технологій інтелектуального аналізу даних (див. рис. 1). Цей метод базується на інтеграції структурованих, неструктурованих і напівструктурованих даних. Формалізація задачі кластерного аналізу виконується через трійку $Z_l = (I_j, D_i, A_k)$, де $I_j = \{I_1, I_2, I_3, I_4\}$ – включає бізнес, соціальну, логістичну та інформаційну інфраструктури; D_i – тип даних (1); $A_k = \{k - means, DBSCAN, GMM, PCA\}$ – обраний метод аналізу. Функція $g(Z_l)$ (2) визначає оптимальне управлінське рішення, яке може включати вибір соціальних груп для розподілу ресурсів, ідентифікацію підприємств з високим потенціалом чи оптимізацію логістичних маршрутів.



Рис. 9. Кроки методу кластерного аналізу соціально-економічних даних ТГ

На основі методу кластерного аналізу соціально-економічних даних ТГ у наступних розділах розкрито інтелектуальні підходи до формування споживчого кошика [24] та пошуку альтернативних постачальників [19]. Використання цього підходу сприяє покращенню якості прийняття управлінських рішень та підтримці сталого розвитку ТГ.

Для реалізації інтелектуального методу формування споживчого кошика було використано мову програмування Python та датасет Customer Personality Analysis з платформи Kaggle, що містить 2240 записів з різними параметрами клієнтів. Після нормалізації даних шляхом розподілу витрат на 24 місяці, застосовано метод кластеризації K–Means для ідентифікації трьох основних кластерів, визначених методом ліктя. Результати кластеризації показали, що кластер 2 охоплює 27 % клієнтів з середнім доходом та витратами: 25 дол. США – на рибні продукти, 100 дол. США – на м'ясо, 20 дол. США – на солодощі, 400 дол. США – на вино та 50 дол. США – на технічні товари (рис. 10). Цей кластер виявився ключовим для формування споживчого кошика, оскільки він включає найбільшу кількість клієнтів з різними віковими категоріями, що дозволяє забезпечити збалансовані рекомендації для різних груп населення.

Впровадження зазначеного методу суттєво підвищило ефективність процесу формування споживчого кошика: час виконання завдань скоротився з 2–60 днів до 1–3 годин, а загальна ефективність системи досягла приблизно 98,75 %. Крім того, інтеграція додаткових алгоритмів кластеризації та механізмів адаптації до нових даних

дозволила підвищити коефіцієнт силуету на 1,2 %, що свідчить про покращення розділення між кластерами.

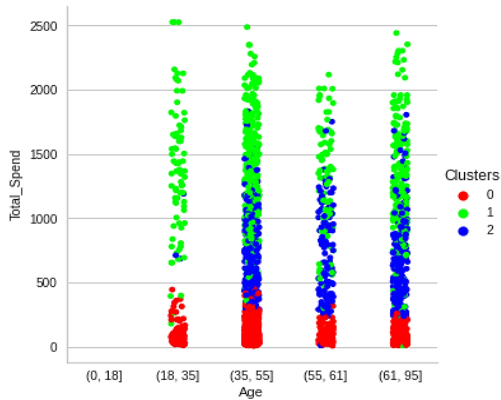


Рис. 10. Візуалізація кластерів: віковий розподіл

Метод пошуку альтернативних постачальників був реалізований на прикладі бізнесу, що спеціалізується на ремонті автомобілів, з використанням мови Python та алгоритму кластеризації K–Means. Для апробації методу зібрано дані про потенційних постачальників з різних джерел, включаючи соціальні медіа та вебсайти, і збережено їх в OLAP базі даних. Застосування методу ліктя дозволило визначити оптимальну кількість кластерів, де для основного кластера Silhouette Score склав 0,78. Вибір найкращого постачальника проводився за ключовими критеріями, такими як ціна, якість, репутація та швидкість доставки. Наприклад, постачальник № 424 був обраний як найкращий кандидат з низькою ціною (10 балів), високою якістю (5 балів), швидкою доставкою (3 дні) та доброю репутацією (5 балів) (рис. 11).

ID	Pricing	Quality	Availability	Production Capacity	Reputation and History	Feedback	Cluster	Selected_Cluster
804	10	3	5	1	5	5	1	2
797	10	3	6	1	1	4	1	2
779	10	3	6	0	2	4	1	2
369	10	3	4	1	3	4	1	2
424	10	5	3	1	5	5	1	2
440	10	2	5	1	1	4	1	2
550	10	2	4	0	4	5	1	2
117	11	4	5	1	3	4	1	2
696	11	5	5	0	3	4	1	2
675	11	4	5	1	4	5	1	2

Рис. 11. Перші 10 значень відсортованих за ціною

Ефективність системи підтверджується кількісними результатами. Час виконання завдань з використанням системи скоротився до 1–5 годин порівняно з 7–60 днями без неї. Загальна ефективність пошуку альтернативних постачальників досягла 99,61 %,

що свідчить про значну оптимізацію процесу. Запропонований підхід забезпечує швидке та точне прийняття рішень, дозволяючи бізнесу ефективно адаптуватися до змін ринку та забезпечувати безперервність операцій навіть у кризових ситуаціях.

Отже, розроблено та впроваджено метод кластерного аналізу соціально-економічних даних ТГ, що дозволяє ефективно підтримувати управлінські рішення через ідентифікацію схожих груп об'єктів. Запропоновані методи формування споживчого кошика та пошуку альтернативних постачальників продемонстрували високу ефективність, скоротивши часові витрати з десятків днів до кількох годин та підвищивши точність прийняття рішень до 99,61%. Результати досліджень підтвердили придатність кластерного аналізу для забезпечення сталого розвитку, оптимізації ресурсів та адаптації до змінних умов у соціально-економічному середовищі територіальних громад.

У шостому розділі представлено метод прогнозування соціально-економічних даних ТГ (рис. 12) [3], який ґрунтується на узагальненому принципі синтезу інформаційних технологій інтелектуального аналізу даних (див. рис. 1). Метод охоплює чотири ключові об'єкти інфраструктури територіальних громад: бізнес-, соціальну, логістичну та інформаційну інфраструктури (I_j), які забезпечують комплексний підхід до збору та аналізу даних. Використання різних типів даних (структуровані, неструктуровані та напівструктуровані) дозволяє глибоко досліджувати соціально-економічні процеси та тенденції.

Ключовим аспектом цього підходу є формалізація задач прогнозування як трійки $Z_l = (I_j, D_i, A_k)$, яка формалізує кожне завдання прогнозування, що включає об'єкт інфраструктури, тип даних (1) та метод аналізу. Ця формула забезпечує гнучкість та адаптивність методу ($A_k = \{LSTM, ARIMA, RF, XGBoost\}$), дозволяючи вибирати найбільш відповідні комбінації для конкретних задач. Результати аналізу цих завдань трансформуються в управлінські рішення через функцію $R_p = g(Z_l)$, що забезпечує підтримку соціально-економічної стабільності та оптимізацію управління ресурсами громад. Таким чином, розроблений метод демонструє високу ефективність та практичну значущість у стратегічному плануванні та прийнятті рішень на рівні територіальних громад.



Рис. 12. Кроки методу прогнозування соціально-економічних даних ТГ

На основі методу прогнозування соціально-економічних даних ТГ у наступних підрозділах деталізовано інтелектуальні підходи до прогнозування попиту на

продукцію [30] та вибору оптимальної локації для старту бізнесу в ТГ [11, 35]. Розроблений метод демонструє значний потенціал для підтримки стратегічного планування та прийняття управлінських рішень у громадах, забезпечуючи адаптивність і точність прогнозування. Зазначений метод успішно реалізовано у низці наукових робіт, зокрема у [40, 46], що підкреслює його актуальність та практичну значущість.

У рамках апробації методу прогнозування соціально-економічних даних було реалізовано проєкт «СмартМед», спрямований на розробку інноваційних медичних рішень для смарт-міст. Для прогнозування ключових параметрів, таких як продажі парацетамолу, відгуки про сервіс та кількість послуг, було застосовано різні моделі машинного навчання. Модель HistGradientBoostingRegressor показала найнижчі значення показників RMSE та MAE для більшості параметрів, що свідчить про її високу точність. Наприклад, для продажів парацетамолу RMSE становив 10,2, а MAE – 8,6, що значно перевершує результати інших моделей, таких як XGBoost та CatBoost.

Особливу увагу приділено прогнозуванню параметра "number of services" за допомогою моделі XGBoost, яка досягла коефіцієнта детермінації $R^2 = 0,98$, що свідчить про надзвичайно високу відповідність прогнозів фактичним даним. На рис. 13 продемонстрована модель інтелектуального прогнозування попиту, де значення RMSE склали приблизно 10,35, а MAE – 8,97, підтверджуючи точність прогнозів. Крім того, у кризових умовах система забезпечила оптимізацію процесу прогнозування з ефективністю близько 95,83 %, що дозволяє підприємствам оперативно реагувати на зміни ринку та ефективно управляти ресурсами. Ці результати підкреслюють практичну цінність запропонованого методу для управління соціально-економічними процесами в територіальних громадах.

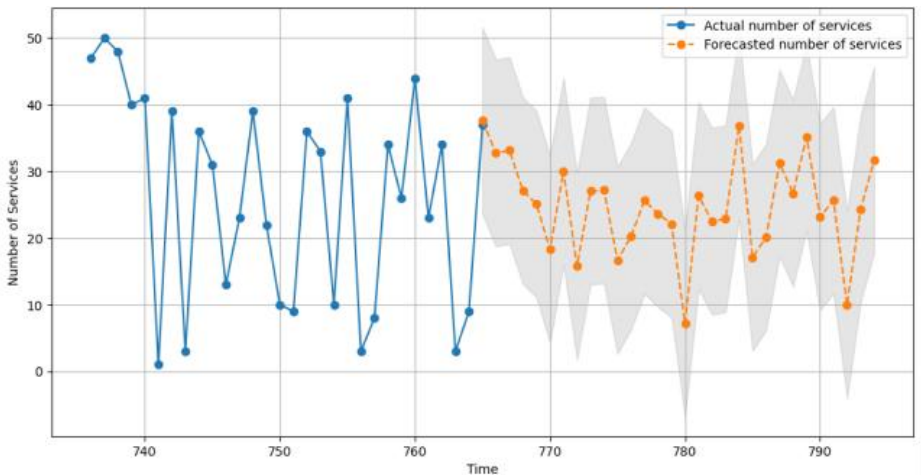


Рис. 13. Інтелектуальне прогнозування попиту

Метод вибору локації для старту бізнесу в територіальних громадах було апробовано на основі аналізу потоку відвідувачів з використанням відеоданих з

YouTube. За допомогою нейронних мереж та бібліотек Python було проаналізовано понад 29 хвилин відео, у результаті чого виявлено 290 осіб. Найбільша група – особи віком 25–32 роки (93 людини), а за гендерним розподілом жінки становили більшість (168 осіб). Зібрані дані були структуровані в таблицю для подальшого аналізу. Після оцінки демографічного складу, для кожного запису було визначено рекомендовані бізнес-категорії, зокрема "Books" (158 випадків), "Technology" (118) та "Fashion" (105). Ці результати підтвердили доцільність відкриття бізнесу, орієнтованого на ці категорії.

Прогнозування потоку відвідувачів здійснювалося за допомогою моделі SARIMA (рис. 14). Аналіз показав поступове зростання кількості відвідувачів протягом перших п'яти годин із піками на шостій (20 % від загального обсягу) та сьомій (30 %) годинах. Значення середньої абсолютної похибки (MAE) склало 9,8, а середньоквадратичної похибки (RMSE) – 12,3, що свідчить про високу точність прогнозів. Такий підхід дозволив ідентифікувати оптимальні часові інтервали для концентрації ресурсів і зниження витрат у періоди низької активності. Загалом, метод досягає ефективності 98,75 %, суттєво скорочуючи час аналізу та підвищуючи точність прийняття рішень щодо вибору бізнес-локацій.

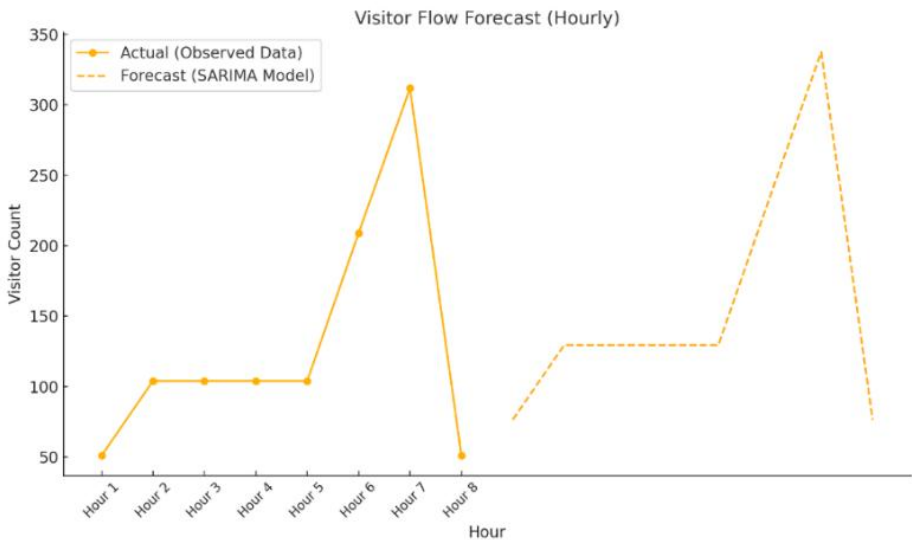


Рис. 14. Прогноз потоку відвідувачів за допомогою SARIMA

Отже, розроблено метод прогнозування соціально-економічних даних територіальних громад, який інтегрує різні типи даних та використовує сучасні методи інтелектуального аналізу. Експериментальні результати демонструють високу точність прогнозів попиту та вибору локацій, з показниками ефективності до 98,75 %, що підтверджується низькими значеннями RMSE та MAE. Запропоновані методи довели свою практичну цінність, сприяючи покращенню управлінських рішень та забезпеченню соціально-економічної стабільності територіальних громад.

У цьому розділі представлено метод гібридного аналізу соціально-економічних даних ТГ [3] (рис. 15), який ґрунтується на узагальненому принципі синтезу інформаційних технологій інтелектуального аналізу даних (див. рис. 1). Гібридний аналіз передбачає два рівні обробки даних. На першому рівні (A^I) застосовуються базові методи класифікації чи кластеризації, представлені множиною:

$$A^I = \{k - \text{means}, \text{DBSCAN}, \text{SVM}, \text{RF}\}. \quad (11)$$

Ці методи дозволяють розподіляти об'єкти на кластери, враховувати щільність даних, класифікувати структуровані дані та аналізувати як кількісні, так і текстові дані.

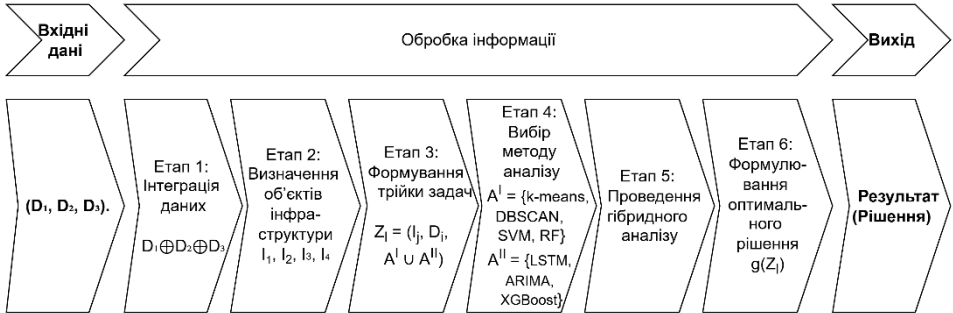


Рис. 15. Кроки методу гібридного аналізу соціально-економічних даних ТГ

Другий рівень (A^{II}) фокусується на детальному прогнозуванні з використанням методів:

$$A^{II} = \{\text{LSTM}, \text{ARIMA}, \text{XGBoost}\}, \quad (12)$$

які забезпечують аналіз часових рядів, прогнозування трендів та точне прогнозування обсягів чи категорій даних. Така дворівнева архітектура дозволяє досягти високої гнучкості та адаптивності методології, забезпечуючи комплексний підхід до розв'язання управлінських задач.

Задачі аналізу формуються як трійки $Z_l = (I_j, D_i, A_k)$, де I_j належить до множини інфраструктурних об'єктів, D_i – до типів даних, а A_k – до набору методів аналізу $A^I \cup A^{II}$. Результати аналізу визначаються функцією $R_p = g(Z_l)$, яка формує управлінські рішення на основі проведеного аналізу, такі як класифікація регіонів за рівнем економічної стабільності чи прогнозування потреб громади у ключових ресурсах.

На основі методу гібридного аналізу соціально-економічних даних ТГ детально розглянуто інтелектуальні методи для класифікації рівня техногенних катастроф [6, 16] та прогнозування обсягів відходів громади [18]. Запропонований підхід забезпечує всебічний аналіз і прийняття стратегічних рішень у кризових ситуаціях, а також дозволяє ефективно вирішувати екологічні та управлінські задачі. Розроблений метод успішно реалізовано у низці наукових робіт, зокрема [13, 15, 19], що підтверджує його актуальність і практичну значущість для соціально-економічного розвитку ТГ.

Гібридний метод для класифікації рівня техногенних катастроф ефективно використовує два типи даних: кількісні показники та текстові описи. На основі записів про аварії на 12 підприємствах, метод розділяє дані на структуровані (кількісні) та

неструктуровані (текстові), забезпечуючи двоетапний аналіз. На першому рівні кількісні дані проходять класифікацію за допомогою методів ансамблевого навчання, таких як Gradient Boost і CatBoost, які забезпечили точність до 92 % і 98 %, відповідно (табл. 3). Текстові описи, зібрані у вигляді деталізованих звітів про інциденти, обробляються для визначення рівня потенційної небезпеки з подальшою класифікацією.

Таблиця 3

Результати моделюванні ансамблевими методами МН

Модель	Точність
AdaBoost	0,595376
Gradient Boost	0,924855
XGBoost	0,901734
CatBoost	0,988439
LGBost	0,791908

У процесі моделювання на тестовій вибірці метод CatBoost досяг середньої точності прогнозування рівня потенційної небезпеки 81 % для вибірки з 382 значень. Особливо висока точність спостерігалася для класифікації рівнів небезпеки V та I, де відповідно були досягнуті показники 0,97 та 0,94.

Крім того, клас "Industry Sector" було класифіковано з точністю 99 % для 282 випадків, а клас "Employee or Third Party" – з абсолютною точністю 100 % для 28 випадків, підтверджуючи високу ефективність моделі в розпізнаванні різних категорій даних. Ці результати демонструють здатність запропонованого методу точно класифікувати рівні техногенних катастроф, забезпечуючи надійний інструмент для управлінських рішень у кризових ситуаціях.

Гібридний підхід дозволив досягти високої адаптивності завдяки інтеграції кількісного та текстового аналізу. Для прогнозування потенційного рівня небезпеки модель CatBoost показала найкращі результати: середня точність на тестовій вибірці склала 81 %, а F1-міра досягла 0,88. Паралельна обробка кількісних і текстових даних забезпечила всебічний аналіз аварійних ситуацій, що підтверджено точністю передбачення ризиків для текстових описів з сектору гірничодобувної промисловості. Висока ефективність методів дозволила автоматизувати процес оцінки ризиків техногенних катастроф, забезпечуючи оперативне прийняття управлінських рішень.

Гібридний метод прогнозування обсягів відходів громади починається з кластеризації даних, що дозволяє сегментувати відходи за типами та ступенем переробки. Для цього використовуються різні методи кластеризації, такі як K-Means, DBSCAN, Birch та Spectral Clustering. Результати кластеризації відображені на рис. 16, де видно, що DBSCAN та Birch ефективно групують схожі типи відходів, дозволяючи виокремлювати кластери з найбільш подібними характеристиками. Наприклад, метод K-Means показав високу стабільність для кластерів з номерами 0 та 1, які характеризуються значною кількістю паперових і пластикових відходів, що генеруються в громадах. Завдяки цьому, адміністрації громади легше визначати групи матеріалів для більш ефективного управління ресурсами та переробкою.

Після кластеризації проводиться прогнозування обсягів відходів для кожного кластера за допомогою моделей ARIMA, DNN та XGBoost. Модель XGBoost показала

найкращі результати з точністю прогнозування до 98 %. Наприклад, для кластерів, створених методами DBSCAN, Birch та Spectral Clustering, прогнозування кількості перероблених відходів продемонструвало мінімальні відхилення за метриками MAE та MSE, що свідчить про високу ефективність підходу. Важливо, що впровадження кластеризації перед прогнозуванням дозволило оптимізувати моделі для кожної групи відходів, підвищуючи точність аналізу.

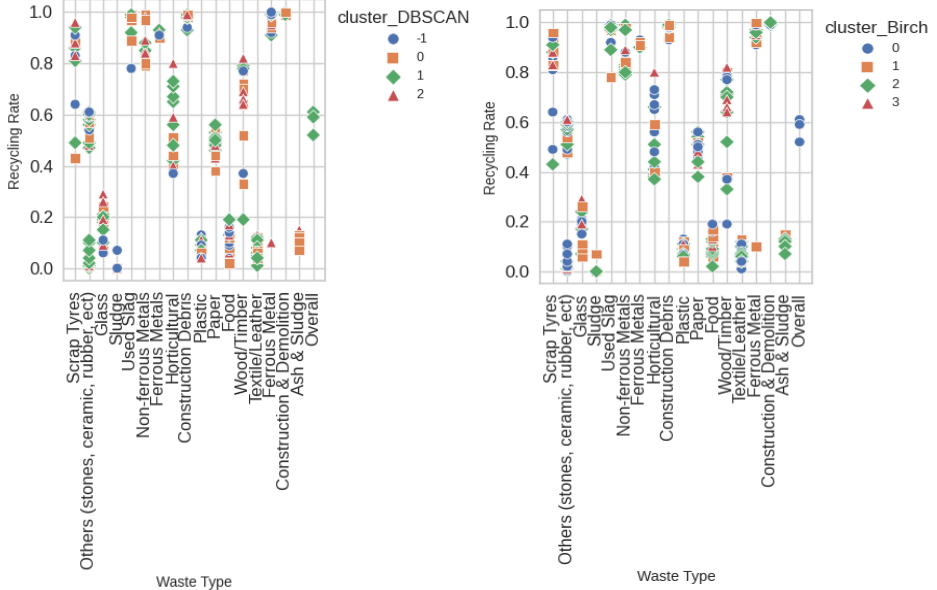


Рис. 16. Кластеризація відходів за типом та ступенем переробки

Гібридний метод забезпечує комплексний підхід, інтегруючи кластеризацію та прогнозування для вдосконалення управління відходами громади. Наприклад, DBSCAN і Birch показали не лише точність у сегментації даних, а й сприяли побудові більш точних прогнозів за допомогою XGBoost, досягнувши 97 % середньої точності для тестових даних. Цей підхід дозволив скоротити час аналізу з 5–10 днів до 1–2 годин, що свідчить про значне покращення у швидкості та ефективності управління відходами. Результати досліджень вказують, що використання кластеризації для сегментації даних та прогнозування у межах кластерів є критично важливим для забезпечення сталого розвитку громади і зменшення впливу на довкілля.

Запропонований гібридний метод аналізу соціально-економічних даних забезпечує ефективну інтеграцію різнорідних даних та високу точність прогнозів, що сприяє оптимізації управлінських рішень у територіальних громадах. Використання методів кластеризації та прогнозування дозволило значно скоротити час аналізу і підвищити ефективність управління ресурсами та безпекою. Отримані результати підтверджують практичну цінність гібридного підходу для забезпечення сталого розвитку та екологічної стабільності громад, що свідчить про його доцільність впровадження у практику управління.

ОСНОВНІ РЕЗУЛЬТАТИ ТА ВИСНОВКИ

У цьому дослідженні розроблено інноваційну інформаційну технологію інтелектуального аналізу соціально-економічних даних ТГ, яка забезпечує підвищення соціальної стійкості та ефективності управління ресурсами. Запропоновані методи та моделі дозволяють інтегрувати різноманітні дані, враховувати динамічність соціально-економічних процесів і приймати обґрунтовані управлінські рішення в режимі реального часу.

Отримано такі основні наукові та практичні результати:

1. Проведено аналіз сучасного стану інтелектуального аналізу соціально-економічних даних для ТГ, яке виявило суттєві протиріччя та обмеженість існуючих методів щодо забезпечення точного прогнозування, адаптивності та інтеграції даних з різних джерел у режимі реального часу. Це підтвердило актуальність розроблення інтелектуальних технологій для ефективного управління та прийняття рішень у динамічних умовах.

2. Вперше запропоновано узагальнені принципи синтезу інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ, які забезпечують інтеграцію структурованих, неструктурованих і напівструктурованих даних та враховують динамічність соціально-економічних процесів. Це дозволило підвищити соціальну стійкість інфраструктури громад за рахунок адаптивного вибору методів аналізу та формування управлінських рішень, скоротивши час на обробку даних та прийняття рішень з днів до годин, що підтверджено експертною ефективністю ~99,72 %.

3. Розроблено методологію аналізу соціально-економічних даних ТГ, що включає сукупність методів класифікаційного, кластерного та гібридного аналізу, а також прогнозування. Особливістю методології є адаптивність до нестационарних даних і можливість роботи в режимі реального часу, що підвищило точність прогнозування, класифікації та прозорість управлінських рішень.

4. Розроблено новий метод формування навчальної вибірки для нестационарних даних на основі RFM-аналізу та кластерного аналізу. Це дозволило більш точно відбирати інформативні зразки, порівнюючи їх з аналогами, що підтверджено експериментальними результатами з точністю класифікації до 99,1 %.

5. Вперше розроблено метод ансамблевих адаптивних прогнозних моделей для багатовимірного аналізу, особливістю якого є оцінювання на «ковзному вікні» та використання метамоделей першого і другого рівнів на основі оптимізації функції Лагранжа. Це забезпечило підвищення точності прогнозів багатовимірних нестационарних процесів до 98 %, що критично для прийняття рішень у реальному часі.

6. Удосконалено метод класифікаційного аналізу кількісних соціально-економічних даних, який забезпечує інтеграцію різноманітних даних. Це дозволило підвищити точність класифікації кількісних показників для реалізації управлінських рішень та скоротити час аналізу з днів до годин. Експертні оцінки показали ефективність методу ~99,72 % у задачі виявлення фіктивних підприємств, що підтверджує значне зменшення часу на ідентифікацію ризиків.

7. Удосконалено метод класифікаційного аналізу текстових соціально-економічних даних з інтеграцією текстових джерел різної структури та застосуванням технологій обробки природної мови. Це дозволило зменшити час виявлення закономірностей з 1–5 днів до 1–2 годин, підвищивши ефективність виявлення закономірностей для реалізації управлінських рішень, що підтверджено експертними оцінками ефективності ~99,17 %.

8. Удосконалено метод кластерного аналізу соціально-економічних даних, який забезпечує ідентифікацію груп об'єктів з подібними характеристиками. Це дозволило підвищити точність розподілу ресурсів для реалізації управлінських рішень, з ефективністю методу ~99,61 % за експертними оцінками.

9. Удосконалено метод прогнозування соціально-економічних даних, який інтегрує різні типи даних та застосовує адаптивні методи інтелектуального аналізу. Це скоротило час прогнозування з днів до годин, підвищивши ефективність процесу на ~98,75 %, що дозволяє швидко реагувати на зміни попиту та інших важливих показників.

10. Удосконалено метод гібридного аналізу соціально-економічних даних, який враховує як кількісні, так і якісні показники через інтеграцію різнорідних даних та багаторівневий підхід до аналізу. Це зменшило час прийняття комплексних рішень з тижнів до годин, підвищивши гнучкість та адаптивність процесу на ~99,72 % за експертними оцінками.

11. Реалізовано розроблену інформаційну технологію інтелектуального аналізу соціально-економічних даних ТГ у вигляді модулів інформаційної системи та проведено розрахунок її ефективності. Впровадження технології підтвердило теоретичні результати на практиці, забезпечивши підвищення точності прогнозів до 98 %, а ефективність кластеризації та класифікації даних до 95 %. Розрахунок ефективності продемонстрував скорочення часових витрат на аналіз даних з 5–30 днів до 1–120 годин. Ці кількісні показники підтверджують практичну цінність розробленої технології для управління в ТГ, забезпечуючи більш швидке та точне прийняття управлінських рішень та оптимізацію використання ресурсів.

Результати дисертаційної роботи впроваджено в освітній процес Західноукраїнського національного університету, зокрема для підготовки здобувачів спеціальності «Комп'ютерні науки» за навчальними курсами «Інтелектуальний аналіз даних», «Машинне навчання», «Методи та засоби штучного інтелекту» та іншими. Також результати використано у науково-дослідних проєктах Erasmus+: «WORK4CE» і «My Farm», та Creative Europe «AURA», а також у державних наукових роботах, зокрема, TruScanAI, «Методи та засоби структурно-статистичної ідентифікації», «Інтелектуальні методи соціально-економічного розвитку ТГ» і в рамках перспективного плану розвитку наукового напрямку «Технічні науки» ЗУНУ. Крім того, результати дисертаційної роботи апробовано в ІТ-компанії «МагнетікВан» (м. Тернопіль), Тернопільській обласній державній адміністрації та Великоберезовицькій ТГ.

СПИСОК ПРАЦЬ, ОПУБЛІКОВАНИХ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових виданнях, включених до Переліку наукових фахових видань України:

1. Лип'яніна-Гончаренко, Х. (2024). Узагальнений принцип синтезу інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ. *Measuring and Computing Devices in Technological Processes*, (1), 359–367. <https://doi.org/10.31891/2219-9365-2024-77-48>
2. Лип'яніна-Гончаренко, Х. (2024). Методологія аналізу соціально-економічних даних територіальних громад. *Herald of Khmelnytskyi National University. Technical Sciences*, 337(3(2)), 446–450. <https://doi.org/10.31891/2307-5732-2024-337-3-67>
3. Лип'яніна-Гончаренко, Х. (2024). Методи аналізу соціально-економічних даних територіальних громад для адаптивного управління ресурсами. *Computer Systems and Information Technologies*, (3), 92–97. <https://doi.org/10.31891/csit-2024-3-12>
4. Лип'яніна-Гончаренко, Х. В. (2023). Метод формування навчальної вибірки для масивів даних на основі машинного навчання. *Наукові записки НаУКМА. Комп'ютерні науки*, (6), 30–35. <https://doi.org/10.18523/2617-3808.2023.6.30-35>
5. Лип'яніна-Гончаренко, Х., & Юрків, Х. (2023). Методи бустингового машинного навчання для нестационарних часових рядів. *Measuring and Computing Devices in Technological Processes*, (3), 19–30. <https://doi.org/10.31891/2219-9365-2023-75-2>
6. Лип'яніна-Гончаренко, Х., Комар, М., Юрків, Х., & Лук'янчук, В. (2023). Концептуальна модель інтелектуальної оцінки наслідків техногенних катастроф. *Вісник Хмельницького національного університету. Серія: Технічні науки*, 329(6), 230–237. <https://doi.org/10.31891/2307-5732-2023-329-6-230-237>
7. Комар, М., Лип'яніна-Гончаренко, Х., Кіт, І., Мадараш, Р., & Юрків, Х. (2023). Інтелектуальний метод виявлення джерел мультилінгвальної дезінформації. *Measuring and Computing Devices in Technological Processes*, (2), 221–230. <https://doi.org/10.31891/2219-9365-2023-74-31>
8. Лип'яніна-Гончаренко, Х. В. (2022). Інтелектуальний метод формування людських ресурсів на короткостроковий проєкт. *Вісник Черкаського державного технологічного університету*, (3), 49–58. <https://doi.org/10.24025/2306-4412.3.2022.259775>
9. Лип'яніна-Гончаренко, Х., Комар, М., Саченко, А., & Лендюк, Т. (2022). Оцінка інвестиційних ризиків віртуальної ІТ-компанії на основі машинного навчання. *Measuring and Computing Devices in Technological Processes*, (3), 45–60. <https://doi.org/10.31891/2219-9365-2022-71-3-6>
10. Лип'яніна-Гончаренко, Х. В., Комар, М. П., Саченко, А. О., & Лендюк, Т. В. (2022). Метод виявлення фіктивних підприємств на підставі Гаусового наївного класифікатора Байєса. *Scientific Bulletin of UNFU*, 32(5), 92–96. <https://doi.org/10.36930/40320513>
11. Лип'яніна-Гончаренко, Х. (2023). Інтелектуальний метод вибору локації для старту бізнесу в розумному місті. *Системні технології*, 4(147), 132–140. <https://doi.org/10.34185/1562-9945-4-147-2023-12>

12. Лип'яніна-Гончаренко, Х., Комар, М., Саченко, А., & Лендюк, Т. (2022). Метод формування контексту реклами та цільової аудиторії на основі навчання асоціативних правил. *Вісник Хмельницького національного університету. Серія: Технічні науки*, 313(5), 279–287. <https://doi.org/10.31891/2307-5732-2022-313-5-279-287>

13. Лип'яніна-Гончаренко, Х. В. (2023). Метод генерування рекламного зображення на основі ключових слів. *Телекомунікаційні та інформаційні технології*, 81(4). <https://doi.org/10.31673/2412-4338.2023.043745>

14. Лип'яніна-Гончаренко, Х., & Кіт, І. (2023). Метод удосконалення рекламних текстів на основі генеративних моделей. *Вісник Приазовського державного технічного університету. Серія: Технічні науки*, (46), 6–13. <https://doi.org/10.31498/2225-6733.46.2023.288087>

15. Лип'яніна-Гончаренко, Х. В. (2023). Метод генерування рекламного зображення на основі відеопотоку. *Науковий вісник Ужгородського університету. Серія: Математика і інформатика*, 43(2), 130–135. [https://doi.org/10.24144/2616-7700.2023.43\(2\).130-135](https://doi.org/10.24144/2616-7700.2023.43(2).130-135)

Статті у періодичних виданнях, включених до категорії «А» Переліку наукових фахових видань України, або у закордонних виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus:

16. Lipianina-Honcharenko, K., Wolff, C., Sachenko, A., Kit, I., & Zahorodnia, D. (2023). Intelligent method for classifying the level of anthropogenic disasters. *Big Data and Cognitive Computing*, 7(3), 157. <https://doi.org/10.3390/bdcc7030157> (Індексована в наукометричній базі Web of Science та Scopus, що відноситься до Q2 відповідно до SCImago Journal & Country Rank)

17. Lipianina-Honcharenko, K., Wolff, C., Sachenko, A., Desyatnyuk, O., Sachenko, S., & Kit, I. (2023). Intelligent information system for product promotion in internet market. *Applied Sciences*, 13(17), 9585. <https://doi.org/10.3390/app13179585> (Індексована в наукометричній базі Web of Science та Scopus, що відноситься до Q2 відповідно до SCImago Journal & Country Rank)

18. Lipianina-Honcharenko, K., Komar, M., Osolinskyi, O., Shymanskyi, V., Havryliuk, M., & Semaniuk, V. (2024). Intelligent waste-volume management method in the smart city concept. *Smart Cities*, 7(1), 78–98. <https://doi.org/10.3390/smartcities7010004> (Індексована в наукометричній базі Web of Science та Scopus, що відноситься до Q1 відповідно до SCImago Journal & Country Rank)

19. Lipianina-Honcharenko, K., Komar, M., Melnyk, N., & Komarnytsky, R. (2024). Sustainable information system for enhancing virtual company resilience through machine learning in smart city socio-economic scenarios. *Economics*, 12(2). <https://doi.org/10.2478/eoik-2024-0022> (Індексована в наукометричній базі Web of Science та Scopus, що відноситься до Q2 відповідно до SCImago Journal & Country Rank)

20. Lipianina-Honcharenko, K., Bodyanskiy, Y., Kustra, N., & Ivasechko, A. (2024). OLTW-TEC: Online learning with sliding windows for text classifier ensembles. *Frontiers in Artificial Intelligence*, 7, 1401126. <https://doi.org/10.3389/frai.2024.1401126> (Індексована в наукометричній базі Scopus та відноситься до Q2 відповідно до SCImago Journal & Country Rank)

21. Bodyanskiy, Y. V., Lipianina-Honcharenko, K. V., & Sachenko, A. O. (2022). Ensemble of adaptive predictors for multivariate nonstationary sequences and its online learning. *Radio Electronics, Computer Science, Control*, (4(67)), 91–97, <https://doi.org/10.15588/1607–3274–2023–4–9> (Індексована в наукометричній базі Web of Science)

22. Lipianina-Honcharenko, K., Bodyanskiy, Y., & Sachenko, A. (2023). Intelligent information system of the city's socio-economic infrastructure. *System Research and Information Technologies*, (3), 108–120. <https://doi.org/10.20535/srit.2308–8893.2023.3.08> (Індексована в наукометричній базі Scopus та відноситься до Q4 відповідно до SCImago Journal & Country Rank)

Монографії (розділи у колективних монографіях):

23. Lipyanina, H., Sachenko, O., Lendyuk, T., Sachenko, A., & Vasyukiv, N. (2020). Intelligent method of forming the HR management short-term project. In *Advances in Intelligent Systems and Computing* (pp. 1045–1055). Cham: Springer International Publishing. https://doi.org/10.1007/978–3–030–63270–0_71

24. Lipianina-Honcharenko, K., Wolff, C., Chyzhovska, Z., Sachenko, A., Lendiuk, T., & Grodskyi, S. (2022). Intelligent method for forming the consumer basket. In *Communications in Computer and Information Science* (pp. 221–231). Cham: Springer International Publishing. https://doi.org/10.1007/978–3–031–16302–9_17 (Індексована в наукометричній базі Scopus)

25. Lipyanina, H., Maksymovych, V., Sachenko, A., Lendyuk, T., Fomenko, A., & Kit, I. (2020). Assessing the investment risk of virtual IT company based on machine learning. In *Communications in Computer and Information Science* (pp. 167–187). Cham: Springer International Publishing. https://doi.org/10.1007/978–3–030–61656–4_11 (Індексована в наукометричній базі Scopus)

26. Lipianina-Honcharenko, K., Lendiuk, T., Sachenko, A., Osolinskyi, O., Zahorodnia, D., & Komar, M. (2022). An intelligent method for forming the advertising content of higher education institutions based on semantic analysis. In *Communications in Computer and Information Science* (pp. 169–182). Cham: Springer International Publishing. https://doi.org/10.1007/978–3–031–14841–5_11 (Індексована в наукометричній базі Scopus)

Статті у матеріалах міжнародних конференцій, які індексуються у наукометричних базах Scopus та Web of Science:

27. Lipyanina-Goncharenko, H., Brych, V., Sachenko, S., Lendyuk, T., Bykovyy, P., & Zahorodnia, D. (2021). Method of forming a training sample for segmentation of tender organizers on a machine learning basis. In *COLINS* (pp. 1843–1852). <https://ceur-ws.org/Vol–2870/paper134.pdf> (Індексована в наукометричній базі Scopus та Web of Science)

28. Lipianina-Honcharenko, K., Kit, I., Zahorodnia, D., & Osolinskyi, O. (2023). Optimization method of advertising texts based on generative models. In *Proceedings of the 3rd International Workshop on Information Technologies: Theoretical and Applied Problems 2023* (Vol. 3628, pp. 126–135). <https://ceur-ws.org/Vol–3628/paper10.pdf> (Індексована в наукометричній базі Scopus)

29. Komar, M., Savchyshyn, R., Lipianina-Honcharenko, K., & Osolinskiy, O. (2023). Intelligent method for counting cars from satellite images. In *IntSol* (pp. 295–303). https://ceur-ws.org/Vol-3538/Short_1.pdf (Індексована в наукометричній базі Scopus)
30. Lipianina-Honcharenko, K., Bodyanskiy, Y., Sachenko, A., Kit, I., Podchasova, T., & Lendiuk, T. (2023). An intelligent method of predicting the demand for goods/services in crisis conditions. In *Proceedings of the 3rd International Workshop of IT-Professionals on Artificial Intelligence (ProfIT AI 2023)* (Vol. 3641, pp. 224–232). <https://ceur-ws.org/Vol-3641/short2.pdf> (Індексована в наукометричній базі Scopus)
31. Lipyanina, H., Sachenko, S., Lendyuk, T., Brych, V., Yatskiv, V., & Osolinskiy, O. (2021). Method of detecting a fictitious company on a machine learning basis. In *Advances in Computer Science for Engineering and Education IV* (pp. 138–146). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-80472-5_12 (Індексована в наукометричній базі Scopus)
32. Krysovatu, A., Lipianina-Honcharenko, H., Sachenko, S., Desyatnyuk, O., Banasik, A., & Lukasevych-Krutnyk, I. (2022). Recognizing the fictitious business entity on a logistic regression basis. In *Proceedings of the 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security* (Vol. 3156, pp. 315–327). <https://ceur-ws.org/Vol-3156/paper15.pdf> (Індексована в наукометричній базі Scopus)
33. Krysovatu, A., Lipyanina-Goncharenko, H., Desyatnyuk, O., & Sachenko, S. (2021). Classification method of fictitious enterprises based on Gaussian Naive Bayes. In *2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)* (pp. 210–216). <https://doi.org/10.1109/CSIT52700.2021.9648584> (Індексована в наукометричній базі Scopus)
34. Krysovatu, A., Lipyanina-Goncharenko, H., Sachenko, S., & Desyatnyuk, O. (2021). Economic crime detection using support vector machine classification. In *Modern Machine Learning Technologies and Data Science Workshop (MoML&T&DS 2021)* (Vol. 2917, pp. 830–840). <https://ceur-ws.org/Vol-2917/paper46.pdf> (Індексована в наукометричній базі Scopus)
35. Lipianina-Honcharenko, K., Sachenko, A., Semaniuk, V., Badasian, A., & Kopania, Ł. (2023). Intelligent method for selecting a business location in a smart city. In *2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. (pp. 1184–1188). <https://doi.org/10.1109/IDAACS58523.2023.10348823> (Індексована в наукометричній базі Scopus)
36. Gramyak, R., Lipyanina-Goncharenko, H., Sachenko, A., Lendyuk, T., & Zahorodnia, D. (2022). Intelligent method for choosing a competitive product based on emotional feedback coloring. In *Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security* (Vol. 2853, pp. 346–357). <https://ceur-ws.org/Vol-2853/paper31.pdf> (Індексована в наукометричній базі Scopus)
37. Lipianina-Honcharenko, K., Lendiuk, T., Sachenko, A., & Wołoszyn, J. (2022). Method of forming the context of advertising and target audience based on associative rule learning. In *Proceedings of the Fifth International Workshop on Computer Modeling and*

Intelligent Systems (CMIS–2022) (Vol. 3137, pp. 98–107). <https://ceur-ws.org/Vol-3137/paper9.pdf> (Індексована в наукометричній базі Scopus)

38. Lipyanina, H., Sachenko, S., Lendyuk, T., & Sachenko, A. (n.d.). Targeting model of HEI video marketing based on a classification tree. In *16th International Conference on ICT in Education, Research, and Industrial Applications. Integration, Harmonization, and Knowledge Transfer (ICTERI 2020)* (Vol. 2732, pp. 487–498). <https://ceur-ws.org/Vol-2732/20200487.pdf> (Індексована в наукометричній базі Scopus)

39. Lipianina-Honcharenko, K., Lukasevych-Krutnyk, I., Butryn-Boka, N., & Sachenko, A. (2022). Intelligent method for identifying fraudulent online stores. In *2021 IEEE 8th International Conference on Problems of Infocommunications, Science, and Technology (PIC S&T)*. <https://doi.org/10.1109/PICST54195.2021.9772195> (Індексована в наукометричній базі Scopus)

40. Koziuk, V., & Lipyanina-Goncharenko, H. (2021). Intelligent method of predicting the discount rate trend. In *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. IEEE. <https://doi.org/10.1109/IDAACS53288.2021.9660835> (Індексована в наукометричній базі Scopus)

41. Lipianina-Honcharenko, K., Sachenko, A., Wolff, C., & Bodyanskiy, Y. (2023). Simulation model for determining the quality of life in Ukrainian cities during the war. In *2023 IEEE European Technology and Engineering Management Summit (E-TEMS)* (pp. 97–101). <https://doi.org/10.1109/E-TEMS57541.2023.10424587> (Індексована в наукометричній базі Scopus)

42. Paryi, V., Lipianina-Honcharenko, K., Brukhanskiy, R., Sachenko, A., Tkachyk, F., & Lendiuk, D. (2023, November). Intelligent verbal interaction methods with non-player characters in metaverse applications. In *2023 IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT)* (pp. 67–71). <https://doi.org/10.1109/AICT61584.2023.10452688> (Індексована в наукометричній базі Scopus)

43. Lipianina-Honcharenko, Kh., Soia, M., Yurkiv, Kh., & Ivasechko, A. (2024). Evaluation of the effectiveness of machine learning methods for detecting disinformation in Ukrainian text data. In *Proceedings of the Seventh International Workshop on Computer Modeling and Intelligent Systems (CMIS–2024)* (pp. 97–109). <https://ceur-ws.org/Vol-3702/paper9.pdf> (Індексована в наукометричній базі Scopus)

44. Sachenko, A., Lendiuk, T., Lipianina-Honcharenko, K., Dobrowolski, M., Boguta, G., & Bytsyura, L. (2024). Method of determining the text sentiment by thematic rubrics. In *COLINS* (pp. 404–414). <https://ceur-ws.org/Vol-3688/paper26.pdf> (Індексована в наукометричній базі Scopus)

45. Lipianina-Honcharenko, K., Melnychuk, A., Yurkiv, K., Hladiy, G., & Telka, M. (2024). Integrated approach to the international aspects of online dispute resolution formation. In *Proceedings of the First International Workshop of Young Scientists on Artificial Intelligence for Sustainable Development* (pp. 88–98). <https://ceur-ws.org/Vol-3716/paper7.pdf> (Індексована в наукометричній базі Scopus)

46. Komar, M., Lipianina-Honcharenko, K., Domanskyi, V., & Melnyk, N. (2024). Regression-based method for real-time solar power plant efficiency forecasting. In *Proceedings of the Modern Machine Learning Technologies Workshop (MoMLeT 2024)* (pp. 235–245). <https://ceur-ws.org/Vol-3711/paper14.pdf> (Індексована в наукометричній базі Scopus)

47. Sachenko, A., Lendiuk, T., Lipianina-Honcharenko, K., Koval, V., Hladyi, G., & Halias, Y. (2024). Evaluation of ensemble machine learning models for movie recommendation systems. In *Proceedings of the Modern Machine Learning Technologies Workshop (MoMLeT 2024)* (pp. 273–286). <https://ceur-ws.org/Vol-3711/paper17.pdf> (Індексована в наукометричній базі Scopus)

48. Lipianina-Honcharenko, K., Maika, N., Sachenko, S., Kopania, L., & Soia, M. (2024). A cyclical approach to legal document analysis: Leveraging AI for strategic policy evaluation. In *CEUR Workshop Proceedings* (Vol. 3736, pp. 201–211). <https://ceur-ws.org/Vol-3736/paper15.pdf> (Індексована в наукометричній базі Scopus)

Публікації, які додатково відображають наукові результати дисертації:

49. Крисоватий, А. І., Лип'яніна-Гончаренко, Х. В., та ін. (2022). *Метод виявлення фіктивних підприємств на основі машинного навчання*. Авторське право на твір № 111509, Україна. Оpubл. 31.03.2022, Бюл. № 69.

АНОТАЦІЇ

Ліп'яніна-Гончаренко Х. В. Теоретичні та прикладні засади інформаційної технології інтелектуального аналізу соціально-економічних даних територіальних громад. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 «Інформаційні технології» (12 – Інформаційні технології). – Хмельницький національний університет, Хмельницький національний університет, Хмельницький, 2024.

Дисертація присвячена вирішенню актуальної науково-прикладної проблеми розроблення теоретичних та прикладних засад інформаційної технології інтелектуального аналізу соціально-економічних даних ТГ. Запропонована технологія інтегрує різномірні джерела інформації, адаптивно враховує динамічні та нестационарні соціально-економічні процеси, забезпечуючи високу точність прогнозування та прозорість результатів аналізу. Розроблена методологія включає класифікаційний, кластерний і гібридний аналіз даних, а також прогностичні моделі, що сприяють формуванню обґрунтованих управлінських рішень, підвищенню соціальної стійкості та забезпеченню сталого розвитку ТГ.

Вирішення окресленої наукової проблеми було досягнуто завдяки проведенню комплексних досліджень, що включали аналіз стану і перспектив розвитку аналізу соціально-економічних даних в інфраструктурі ТГ, розробку узагальненого принципу синтезу інформаційної технології для інтелектуального аналізу цих даних, створення методології аналізу з адаптивними методами класифікації, кластеризації та прогнозування, розробку нових методів формування навчальної вибірки для нестационарних даних на основі RFM та кластерного аналізу, а також методів ансамблевих адаптивних прогностичних моделей для багатовимірного аналізу. Крім

того, були розроблені методи класифікаційного аналізу кількісних та текстових соціально-економічних даних, методи кластерного аналізу, прогнозування та гібридного аналізу соціально-економічних даних. Оцінка ефективності запропонованих методів та інформаційної технології підтвердила їхню практичну цінність, що дозволило успішно інтегрувати розроблені рішення в управлінські процеси ТГ, сприяючи підвищенню соціальної стійкості та забезпеченню сталого розвитку.

Ключові слова: соціально-економічні дані, методологія, інформаційна технологія, інтелектуальний аналіз даних, класифікація, кластеризація, прогнозування, адаптивні методи, сталий розвиток, територіальні громади.

Lypianina-Honcharenko Kh. V. Theoretical and Applied Foundations of Information Technology for the Intelligent Analysis of Socio-Economic Data of Territorial Communities. – Qualifying scientific work on the right of the manuscript.

The thesis (dissertation) for the degree of Doctor of technical sciences for specialty 05.13.06 «Information technology» (12 – Information technology). – Khmelnytskyi National University, Khmelnytskyi National University, Khmelnytskyi, 2025.

The dissertation addresses a pressing scientific and applied problem concerning the development of theoretical and practical foundations for an information technology for intelligent analysis of the socio-economic data of territorial communities. The proposed technology integrates heterogeneous data sources and adaptively considers dynamic and non-stationary socio-economic processes, ensuring high forecasting accuracy and transparency of analysis results. The developed methodology includes classification, clustering, and hybrid data analysis, as well as predictive models that facilitate the formation of well-grounded management decisions, enhance social resilience, and promote sustainable development within territorial communities.

The resolution of the outlined scientific problem was achieved through comprehensive research, which encompassed the analysis of the state and prospects of socio-economic data analysis within the infrastructure of territorial communities, the development of a generalized principle for synthesizing an information technology for intelligent data analysis, and the creation of a methodology that incorporates adaptive methods for classification, clustering, and forecasting. New methods were developed for constructing training datasets for non-stationary data based on RFM and cluster analysis, as well as ensemble-based adaptive predictive models for multidimensional analysis. Additionally, methods for classification analysis of quantitative and textual socio-economic data, clustering, forecasting, and hybrid analysis of socio-economic data were introduced. The evaluation of the proposed methods and information technology demonstrated their practical value, enabling the successful integration of the developed solutions into the management processes of territorial communities, thereby enhancing social resilience and ensuring sustainable development.

Keywords: socio-economic data, methodology, information technology, intelligent data analysis, classification, clustering, forecasting, adaptive methods, sustainable development, territorial communities.

Підписано до друку 08.01.2025. Формат 60*84 1/16. Цифровий друк.
Ум. друк. арк. – 1,9. Обл.-вид. арк. – 2,0
Наклад 30 прим. Зам. № 0603, 2025

Видавець та виготувач:
ФОП Осадца Ю. В.
м. Тернопіль, вул. 15 Квітня, 2Д/10
тел. (097) 988-53-23

*Свідоцтво про внесення суб'єкта видавничої справи до державного
реєстру видавців, виготівників і розповсюджувачів видавничої продукції
серія ТР № 46 від 07 березня 2013 р.*