

УДК 004.942:628.4:504.064

**А. М. Тригуба¹, Н. Я. Коваль², І. А. Тригуба¹,
І. Р. Фірман², В. Ю. Фамуляк¹**

¹ Львівський національний університет ветеринарної
медицини та біотехнологій імені С.З. Ґжицького
вул. Пекарська, 50, 79010 Львів, Україна

² Львівський державний університет безпеки життєдіяльності
вул. Клепарівська, 35, 79007 Львів, Україна
e-mail: trianamik@gmail.com

Метод ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів

Розглянуто метод ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів в умовах багатофакторної невизначеності міського середовища. Особливу увагу приділено використанню інформаційно-ентропійного підходу для оцінювання значущості ознак, які характеризують соціально-демографічні, просторові, інфраструктурні, виробничі й екологічні параметри системи управління органічними відходами. Інструментарій підтримки прийняття рішень розглядається в контексті формування інтегрального індексу екологічного ризику, ранжування факторів впливу та прогнозування ризикових станів на основі моделей машинного навчання. На основі проведеного аналізу сформовано послідовність процедур ентропійного відбору ознак, інтегрування показників у єдиний індекс і прогнозного моделювання для підтримки управлінських рішень у сфері муніципального поводження з органічними відходами.

Ключові слова: муніципальні органічні відходи, екологічний ризик, багатофакторні дані, ентропійний аналіз, інтегральний індекс, машинне навчання, підтримка прийняття рішень, прогнозування.

Вступ і постановка проблеми

Системи управління муніципальними органічними відходами належать до складних багатофакторних предметних областей, що характеризуються високим рівнем невизначеності, просторовою неоднорідністю та значною кількістю взаємопов'язаних факторів різної природи. Формування екологічних ризиків у таких сис-

темах залежить від соціально-демографічних, просторових, інфраструктурних, виробничих, екологічних і часових параметрів, частина яких має кількісне вираження, тоді як інші формують непрямий або відкладений вплив на стан муніципального середовища. Унаслідок цього традиційні підходи до оцінювання ризиків часто не забезпечують достатньої чутливості до реальної динаміки і обсягів утворення органічних відходів, а також не дозволяють своєчасно формувати управлінські рішення в межах окремих міських громад [1, 2].

Сучасні дослідження у сфері прогнозування поведінки з муніципальними відходами дедалі частіше базуються на використанні моделей машинного навчання, здатних враховувати складні нелінійні залежності між факторами генерації відходів, інфраструктурними характеристиками та соціально-економічними умовами [3–5]. Разом із тим більшість існуючих моделей орієнтована переважно на прогнозування загальних обсягів відходів, тоді як задачі інтегрованого оцінювання екологічного ризику залишаються недостатньо формалізованими.

Особливістю муніципального управління органічними відходами є необхідність прийняття рішень за умов неповної, неоднорідної і частково розрізної інформації. У практичних умовах це проявляється під час визначення пріоритетних територій для розвитку інфраструктури, оптимізації логістики збору, оцінювання навантаження на потужності перероблення, прогнозування сезонних змін обсягів органічної маси й оцінювання потенційних екологічних наслідків [6, 7].

У сучасних підходах управління на основі даних (data-driven) важливого значення набуває не лише використання великої кількості ознак, а й визначення їхньої реальної інформативності у формуванні цільового показника. Саме тому останні дослідження активно використовують ентропійні методи оцінювання значущості ознак, які дозволяють формувати більш обґрунтовані структури вхідних даних для прогностичних моделей [8–10]. Подібний підхід забезпечує зменшення надлишковості даних, підвищення інтерпретованості результатів і стійкість моделей до шумових факторів [11–13].

Особливої актуальності це набуває в задачах формування інтегрального індексу екологічного ризику, де необхідно одночасно враховувати просторові характеристики території, соціально-демографічні параметри, екологічний стан, інтенсивність генерації органічних відходів та інфраструктурну спроможність системи управління [14–16].

Наявні дослідження демонструють ефективність використання методів машинного навчання для окремих завдань у сфері муніципального поводження з відходами, однак питання інтеграції ентропійного ранжування ознак, побудови інтегрального індексу ризику та прогностичного аналізу в межах єдиного методу підтримки прийняття рішень залишаються недостатньо розробленими [17, 18].

Таким чином, науково-прикладна задача, що розглядається у статті, полягає в розробленні методу ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів, який забезпечує виявлення найбільш значущих факторів ризику, інтегрування багатфакторних даних у єдиний показник і прогнозування його змін для підтримки обґрунтованих управлінських рішень на рівні міських територіальних громад.

Стан проблеми: огляд наявних підходів

Протягом останніх років у сфері управління муніципальними органічними відходами активно розвиваються підходи, що спрямовані на прогнозування обсягів утворення відходів, оцінювання екологічного навантаження та підтримку управлінських рішень у складних багатофакторних системах. Основною особливістю таких задач є наявність значної кількості взаємопов'язаних соціально-демографічних, просторових, виробничих, інфраструктурних і екологічних чинників, які одночасно впливають на формування ризикових ситуацій у муніципальному середовищі [19, 20].

У сучасних дослідженнях для прогнозування обсягів твердих побутових відходів широко застосовуються моделі машинного навчання, які дозволяють виявляти приховані нелінійні залежності між вхідними параметрами та цільовими показниками. Такі підходи забезпечують вищу точність прогнозування порівняно з традиційними статистичними моделями, особливо в умовах великої кількості змінних і нерівномірності вихідних даних [21, 22].

Окремий напрям сучасних досліджень пов'язаний із визначенням інформативності ознак, що використовуються у прогнозних моделях. У багатофакторних екологічних системах надлишковість вхідних параметрів часто знижує стійкість моделей та ускладнює інтерпретацію результатів. Тому все ширше використовуються методи ранжування ознак за ступенем їхнього внеску у формування результату прогнозу, серед яких особливе місце займає ентропійний підхід [23].

Ентропійний аналіз дозволяє кількісно оцінити рівень інформаційної значущості кожної ознаки на основі її впливу на цільову змінну. Такий підхід дає можливість не лише зменшити розмірність вхідного простору, але й сформуванати більш обґрунтовану структуру інтегрального показника ризику [24].

У сучасних дослідженнях екологічного моделювання дедалі більшого значення набувають моделі, що дозволяють пояснювати вплив окремих факторів на отриманий результат. Це особливо важливо в задачах муніципального управління, де практичне рішення має ґрунтуватися не лише на точності прогнозу, а й на розумінні причин формування ризику [25].

Для інтегрованого оцінювання стану систем поводження з органічними відходами широко використовуються інтегральні показники, що поєднують у собі демографічні характеристики, площу території, віддаленість від екологічно чутливих зон, потужності перероблення, логістичне навантаження та екологічні параметри [26]. Однак у більшості існуючих підходів вагові коефіцієнти таких показників задаються експертно або приймаються фіксованими, що знижує адаптивність моделі до конкретних умов території [27].

У вітчизняних дослідженнях значна увага приділяється формуванню інструментарію підтримки прийняття рішень у складних слабо структурованих предметних областях, де рішення приймаються на основі неповної або неоднорідної інформації [28, 29]. Показано, що поєднання математичних процедур ранжування, багатокритеріального аналізу та прогнозних моделей дозволяє підвищити обґрунтованість рішень у складних системах [30].

Особливостями сучасних підходів є:

- використання великих масивів багатофакторних даних;
- попереднє визначення інформативності ознак;

- побудова інтегральних показників ризику;
- поєднання прогнозування та інтерпретації результатів;
- орієнтація на підтримку прийняття рішень у муніципальному управлінні.

Таким чином, аналіз сучасних підходів показує, що хоча окремі елементи ентропійного аналізу, ранжування ознак і прогностичного моделювання вже активно застосовуються, їхнє поєднання в єдиному методі оцінювання інтегрального індексу екологічного ризику від муніципальних органічних відходів потребує подальшого розвитку.

Мета та постановка завдання дослідження

Метою дослідження є розроблення методу ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів на основі багатофакторних даних для підвищення обґрунтованості управлінських рішень у системі муніципального поводження з відходами.

Актуальність досягнення поставленої мети визначається потребою врахування великої кількості взаємопов'язаних чинників, що характеризують соціально-демографічні особливості території, просторову структуру, інфраструктуру забезпеченість, виробничі потоки органічної маси і екологічний стан середовища. За таких умов існує потреба у поєднанні процедур оцінювання інформативності ознак, інтегруванні показників у єдиний індекс ризику та прогнозуванні його змін у межах єдиної аналітичної послідовності [31–33].

Для досягнення поставленої мети сформульовано такі завдання дослідження:

- 1) обґрунтувати метод ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів на основі визначення інформаційної значущості багатофакторних показників;
- 2) виконати ентропійну підготовку багатофакторного масиву даних, визначити вагові коефіцієнти показників та на основі запропонованого методу реалізувати прогнозування інтегрального індексу екологічного ризику для підтримки управлінських рішень у системі муніципального поводження з органічними відходами.

Метод ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів

Запропонований метод орієнтований на обробку багатофакторних муніципальних даних, визначення інформативності ознак, формування інтегрального індексу екологічного ризику та прогнозування його змін у часово-просторовому середовищі. Метод реалізується як послідовність взаємопов'язаних етапів, де кожний наступний етап використовує результати попереднього (рис. 1).

Етап 1. Формування та структуризація багатофакторного масиву даних. Перший етап запропонованого методу полягає у формуванні впорядкованого багатофакторного масиву вхідних даних, який відображає сукупність кількісних характеристик, що визначають екологічний стан системи поводження з органічними відходами на рівні муніципальних територій. Саме на цьому етапі створюється аналі-

тична основа для подальшого нормування, оцінювання інформативності ознак, інтегрального індексування та машинного прогнозування.

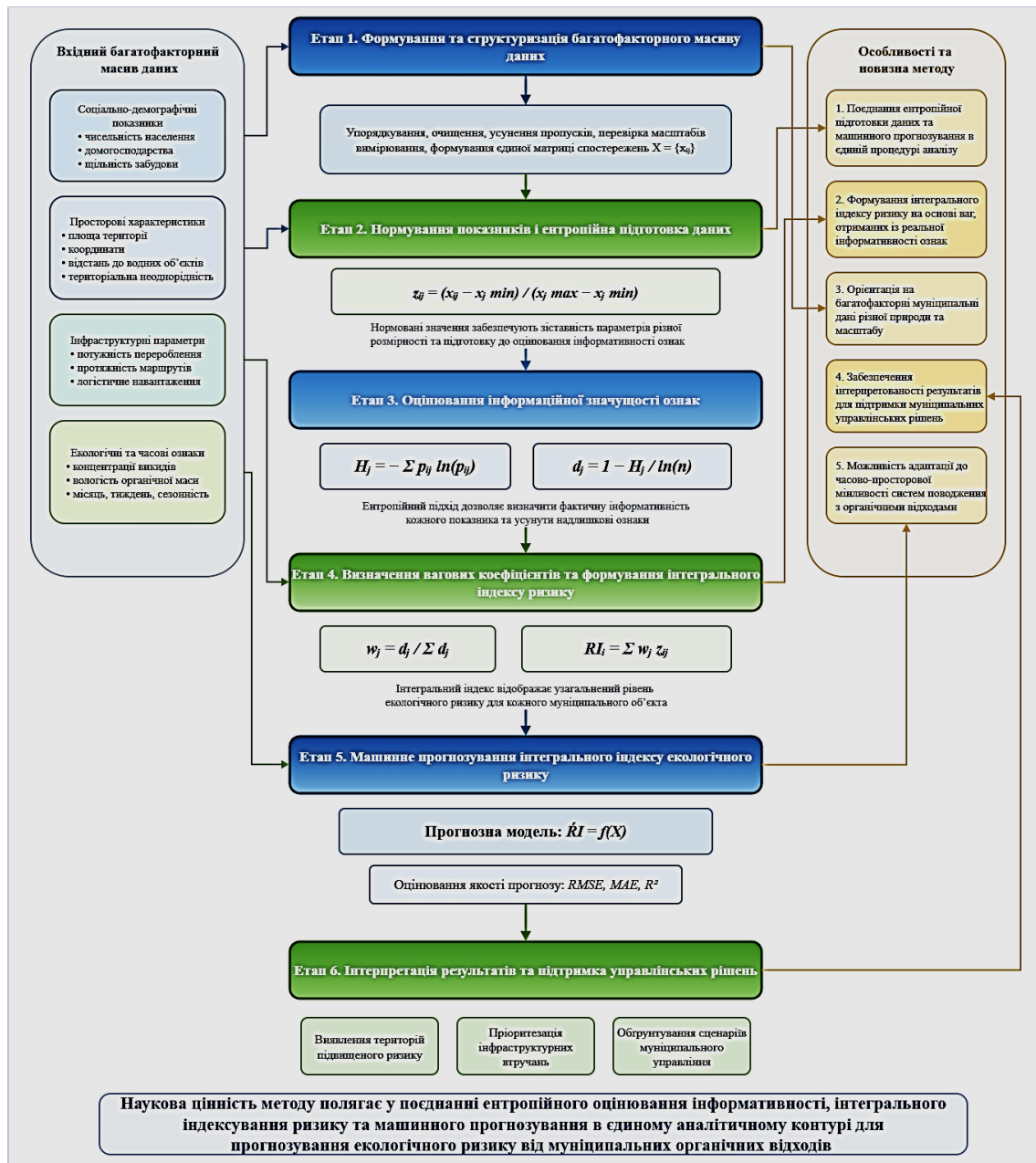


Рис. 1. Схема методу ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів

У структурі методу вхідний масив охоплює кілька функціональних груп показників: соціально-демографічні, просторові, інфраструктурні та екологічні характеристики. До соціально-демографічних належать чисельність населення, кількість домогосподарств, рівень урбанізації. Просторовий блок включає площу території, географічні координати, віддаленість від водних об'єктів і транспортну доступ-

ність. Інфраструктурна група характеризує потужності перероблення, протяжність маршрутів збору відходів, кількість транспортних операцій. Екологічний блок містить показники концентрації метану, вологості органічної маси, температурного режиму, сезонності утворення органічної фракції.

Формально початковий масив спостережень подається у вигляді матриці:

$$X = [x_{ij}], \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (1)$$

де m — кількість муніципальних об'єктів спостереження; n — кількість ознак; x_{ij} — значення j -ї ознаки для i -ї територіальної одиниці.

У сформованій матриці кожен рядок відповідає окремій муніципальній одиниці, а кожен стовпець — конкретному фактору, що потенційно впливає на формування екологічного ризику. На цьому етапі здійснюється перевірка повноти даних, виявлення пропусків, усунення дублювання та логічний контроль меж допустимих значень.

Для забезпечення узгодженості показників проводиться первинна статистична перевірка:

$$\bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij}, \quad (2)$$

де \bar{x}_j — середнє значення кожної ознаки, що дозволяє оцінити загальну структуру розподілу даних перед переходом до нормування.

Додатково визначається варіативність кожної ознаки:

$$\sigma_j = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2}, \quad (3)$$

Отримане стандартне відхилення використовується для виявлення нестабільних або слабоінформативних параметрів, які можуть негативно впливати на подальший розрахунок ентропійних характеристик.

Особливістю цього етапу є те, що структура вхідного масиву формується не лише за принципом доступності даних, а відповідно до логіки причинно-наслідкового впливу показників на екологічний ризик. Це означає, що до моделі включаються лише ті змінні, які здатні відобразити як інтенсивність утворення органічних відходів, так і потенційний рівень екологічного навантаження на територію [34, 35].

Для зручності подальшого аналізу, доцільно подати групування ознак (табл. 1).

Таблиця 1. Структура багатофакторного масиву вхідних даних

Група показників	Приклади ознак	Функціональна роль
Соціально-демографічні	population, households, urban_share	Формування обсягів органічних відходів
Просторові	area km ² , lat, lon, dist to water km	Просторовий контекст ризику
Інфраструктурні	treatment_capacity_tpd, route_km, trips	Характеристика системи збору і перероблення
Екологічні	ch4_ppm, moisture_pct, temperature	Відображення екологічного навантаження

Таким чином, перший етап формує структуровану інформаційну основу всього методу, забезпечуючи перехід від розрізаних статистичних і геопросторових показників до єдиного аналітичного середовища, придатного для подальшого ентропійного оцінювання та прогнозного моделювання. Саме від якості побудови цього масиву залежить точність наступних етапів методу та достовірність інтегрованого індексу екологічного ризику.

Етап 2. Нормування показників і ентропійна підготовка даних. Другий етап методу спрямований на приведення багатофакторного масиву даних до єдиної порівнюваної шкали та підготовку інформаційної основи для подальшого ентропійного оцінювання значущості ознак. Оскільки початкові показники мають різну фізичну природу, одиниці вимірювання та діапазони зміни, без попереднього нормування їх спільне використання у єдиній моделі може призводити до статистичного викривлення результатів. На цьому етапі кожне значення ознаки трансформується у безрозмірний нормований показник, що відображає відносне положення конкретного спостереження в межах сукупності всіх муніципальних об'єктів. Для цього використовується лінійне нормування за мінімальним і максимальним значенням ознаки:

$$z_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}}, \quad (4)$$

де z_{ij} — нормоване значення j -ї ознаки для i -го об'єкта; x_{ij} — початкове значення показника; x_j^{\min} — мінімальне значення ознаки; x_j^{\max} — максимальне значення ознаки.

У результаті такого перетворення всі значення переходять у діапазон від 0 до 1, що забезпечує однаковий масштаб впливу кожного параметра у подальших обчисленнях.

Для показників, зростання яких означає підвищення екологічного ризику (наприклад, концентрація метану, обсяг органічних відходів, навантаження на маршрути), використовується пряме нормування. Для показників, збільшення яких має позитивний вплив на екологічну ситуацію (наприклад, потужність перероблення або віддаленість від чутливих природних зон), застосовується обернене перетворення:

$$z_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}}. \quad (5)$$

Це дозволяє привести всі показники до єдиної логіки: більше значення нормованої ознаки відповідає вищому потенційному екологічному ризику.

Після нормування формується матриця відносних часток, яка використовується для ентропійного аналізу:

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^m z_{ij}}. \quad (6)$$

де p_{ij} — показник, що характеризує частку інформаційного внеску конкретного спостереження в межах ознаки.

Саме ця матриця є основою для подальшого обчислення інформаційної ентропії, оскільки відображає внутрішню структуру розподілу значень кожної змінної. Якщо розподіл значень ознаки рівномірний, її інформативність знижується; якщо ж спостерігається значна варіативність, ознака набуває більшої аналітичної ваги.

Для додаткової перевірки структури нормованих даних визначається середнє нормоване значення (2), а також дисперсія:

$$D_j = \frac{1}{m} \sum_{i=1}^m (z_{ij} - \bar{z}_j)^2. \quad (7)$$

Ці характеристики дозволяють оцінити ступінь концентрації інформації в окремих групах показників перед переходом до етапу ентропійного ранжування. Отримана нормована матриця формує статистично узгоджену основу для третього етапу — оцінювання інформаційної значущості ознак на основі ентропійного підходу. Особливістю цього етапу є те, що нормування не лише усуває масштабні відмінності між показниками, а й створює передумови для кількісного вимірювання інформаційної насиченості кожної змінної. Саме тому ентропійна підготовка даних виступає основною ланкою між первинним багатофакторним описом системи та побудовою інтегрального індексу екологічного ризику

Етап 3. Оцінювання інформаційної значущості ознак. Третій етап методу передбачає кількісне визначення інформативності кожної ознаки шляхом оцінювання ступеня впорядкованості її розподілу у нормованому багатофакторному масиві даних. Саме на цьому етапі визначається, які показники несуть найбільший аналітичний внесок у формування інтегральної оцінки ризику, а які мають меншу розрізняльну здатність.

Основою розрахунку є інформаційна ентропія, яка дозволяє оцінити рівень невизначеності розподілу значень окремої ознаки між муніципальними об'єктами. Якщо значення певного показника близькі між собою, така ознака містить менше корисної інформації для диференціації територій. Якщо ж спостерігається значна неоднорідність, ознака має вищу аналітичну значущість. Ентропія для кожної ознаки визначається за формулою:

$$H_j = - \sum_{i=1}^m p_{ij} \ln(p_{ij}). \quad (8)$$

де H_j — ентропія j -ї ознаки; p_{ij} — нормована частка значення ознаки; m — кількість муніципальних об'єктів.

Для забезпечення порівнюваності ентропійних значень між усіма ознаками використовується нормування ентропії:

$$e_j = \frac{H_j}{\ln(m)}. \quad (8)$$

де e_j — показник, що набуває значень у межах від 0 до 1.

Чим ближче значення e_j до одиниці, тим меншою є інформативність ознаки, оскільки її розподіл наближений до рівномірного. Для визначення реальної інформаційної цінності використовується показник дивергенції:

$$d_j = 1 - e_j. \quad (9)$$

Показник d_j характеризує ступінь відхилення ознаки від стану максимальної невизначеності та фактично відображає її аналітичну значущість у моделі.

Одержані значення дивергенції дозволяють побудувати ранжування всіх змінних за рівнем інформаційного внеску. Ознаки з вищими значеннями d_j надалі формують більший вплив при побудові інтегрального індексу екологічного ризику. Для перевірки розподілу інформативності ознак доцільно формувати впорядкований перелік результатів.

Додатково можна визначити сумарний внесок груп ознак:

$$D = \sum_{j=1}^n d_j. \quad (10)$$

де D — показник, що характеризує загальний інформаційний ресурс багатofакторної системи.

Особливістю третього етапу є те, що оцінювання значущості здійснюється без експертного втручання, лише на основі внутрішньої структури даних. Це підвищує об'єктивність побудови інтегральної моделі та усуває суб'єктивний вплив при визначенні вагових коефіцієнтів. Таким чином, ентропійне оцінювання дозволяє перейти від рівнозначного набору показників до науково обґрунтованої системи пріоритетів, що створює основу для четвертого етапу — визначення вагових коефіцієнтів і формування інтегрального індексу екологічного ризику.

Етап 4. *Визначення вагових коефіцієнтів і формування інтегрального індексу екологічного ризику.* Четвертий етап методу є центральним у загальній структурі дослідження, оскільки саме на цьому етапі відбувається перехід від оцінювання інформативності окремих показників до формування узагальненого числового показника ризику для кожної муніципальної території. Отримані на попередньому етапі значення дивергенції використовуються для визначення вагових коефіцієнтів, які відображають реальний внесок кожної ознаки у формування інтегральної оцінки.

Ваговий коефіцієнт кожної ознаки визначається пропорційно її інформаційній значущості:

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j}. \quad (11)$$

де w_j — ваговий коефіцієнт j -ї ознаки; d_j — значення дивергенції ознаки; n — кількість ознак у моделі.

Сума всіх вагових коефіцієнтів w_j дорівнює одиниці — $\sum_{j=1}^n w_j = 1$. Це забез-

печує математичну збалансованість інтегральної оцінки та дозволяє інтерпретувати кожну вагу як частку впливу відповідного фактора у загальній структурі ризику.

Після визначення ваг формується інтегральний індекс екологічного ризику для кожної муніципальної території:

$$RI_i = \sum_{j=1}^n w_j z_{ij}. \quad (12)$$

де RI_i — інтегральний індекс екологічного ризику для i -го муніципального об'єкта; z_{ij} — нормоване значення j -ї ознаки; w_j — ваговий коефіцієнт ознаки.

Таким чином, кожна територія громади отримує узагальнену кількісну оцінку, яка відображає сукупний рівень екологічного навантаження від муніципальних органічних відходів з урахуванням просторових, інфраструктурних, соціально-демографічних та екологічних характеристик.

Для підвищення інтерпретаційної зручності інтегральний індекс поділений на рівні ризику:

- 1) $RI_i \in [0; 0.33]$ — низький ризик;
- 2) $RI_i \in (0.33; 0.66]$ — середній ризик;
- 3) $RI_i \in (0.66; 1]$ — високий ризик.

Такий підхід дозволяє безпосередньо використовувати результати в управлінських процедурах просторового планування, визначення пріоритетності інфраструктурних втручань і прогнозування зон екологічної напруги. Особливістю цього етапу є те, що інтегральний індекс формується не через експертне вагове оцінювання, а через математично обґрунтовану процедуру, яка спирається на інформаційну структуру самих даних. Це дозволяє мінімізувати суб'єктивність і підвищити відтворюваність результатів. Крім того, сформований індекс стає цільовою змінною для наступного етапу — машинного прогнозування, де на його основі будується модель передбачення екологічного ризику для нових муніципальних умов. Таким чином, четвертий етап завершує побудову інтегральної аналітичної основи методу та формує кількісну платформу для переходу до прогностичного моделювання.

Етап 5. *Прогностичне моделювання інтегрального індексу екологічного ризику.* П'ятий етап методу спрямований на побудову прогностичних моделей, які дозволяють передбачати значення інтегрального індексу ризику для нових територіальних умов або майбутніх часових періодів. На відміну від попередніх етапів, де формувалася аналітична основа оцінювання, тут реалізується машинне навчання як інструмент кількісного прогнозування поведінки ризику в багатофакторному середовищі.

Як цільова змінна використовується інтегральний індекс:

$$y = RI_i, \quad (13)$$

де RI_i — розраховане значення інтегрального екологічного ризику для i -го муніципального об'єкта.

Вхідний вектор прогнозування формується із сукупності найбільш інформативних ознак, визначених на попередньому ентропійному етапі:

$$X = (x_1, x_2, \dots, x_k), \quad (14)$$

де k — кількість ознак, які відібрано після ранжування за інформаційною значущістю.

На цьому етапі було протестовано чотири типи моделей машинного навчання: Ridge Regression, Random Forest, Gradient Boosting та XGBoost. Для кожної моделі виконувалося двоетапне порівняння: базове навчання на повному наборі ознак і

навчання після інтеграції ентропійного ранжування, коли в модель подавалися лише найбільш інформативні фактори.

Лінійна модель Ridge Regression описується залежністю:

$$\hat{y} = \beta_0 + \sum_{j=1}^k \beta_j x_j, \quad (15)$$

де β_j — оцінені параметри моделі.

Для деревоподібних ансамблевих моделей прогноз формується як композиція слабких регресорів:

$$\hat{y} = \sum_{t=1}^T f_t(X), \quad (15)$$

де $f_t(X)$ — окремі дерева рішень у складі ансамблю.

Точність прогнозування оцінювалася за трьома основними метриками: середньою абсолютною похибкою, середньоквадратичною похибкою та коефіцієнтом детермінації.

1. Середня абсолютна похибка:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (16)$$

2. Середньоквадратична похибка:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (17)$$

3. Коефіцієнт детермінації:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad (18)$$

де y_i — фактичне значення інтегрального індексу ризику; \hat{y}_i — прогнозоване значення інтегрального індексу ризику; \bar{y} — середнє значення інтегрального індексу ризику фактичних спостережень.

Важливою особливістю запропонованого методу є можливість його адаптації до нових масивів даних, що накопичуються у процесі функціонування муніципальної системи поводження з органічними відходами. Оскільки інтегральний індекс екологічного ризику формується на основі статистичних залежностей між показниками середовища, їхня інформаційна значущість не є сталою величиною і змінюється залежно від динаміки соціально-економічних, інфраструктурних та екологічних чинників проектного середовища території. У зв'язку з цим після надходження нових даних реалізується процедура періодичного оновлення вагових коефіцієнтів ознак шляхом повторного ентропійного оцінювання їхньої інформативності. На кожному циклі оновлення формується розширена навчальна вибірка, до якої додаються нові часові спостереження, після чого для кожного показника x_i повторно визначається нормалізований ентропійний показник за формулою (8). На основі оновленого ентропійного значення коригується коефіцієнт вагомості показника. Таким чином, ваги автоматично відображають актуальний рівень інформаційної корис-

ності ознак, а набір чинників, що входять до інтегрального індексу ризику, змінюється залежно від появи нових закономірностей у даних.

Після оновлення ентропійних ваг виконується повторний етап відбору інформативних змінних, де до прогнозованої моделі включаються лише ті чинники, для яких значення w_j перевищує встановлений адаптивний поріг відбору. Це дозволяє уникати накопичення слабоінформативних або статистично нестабільних показників у структурі моделі. Далі прогнозна модель проходить повторне навчання на актуалізованій вибірці.

Особливістю п'ятого етапу є те, що прогностична модель не функціонує ізольовано, а інтегрується у попередньо сформовану ентропійну структуру даних. Це означає, що машинне навчання працює вже на інформаційно очищеному просторі ознак, де усунуто слабоінформативні параметри, а головний внесок зосереджений у найбільш значущих факторах. Таким чином, п'ятий етап завершує побудову методу, забезпечуючи перехід від статичної оцінки ризику до його прогнозного відтворення, що відкриває можливість практичного використання результатів у цифрових системах підтримки прийняття рішень у сфері управління муніципальними органічними відходами.

Етап 6. Інтерпретація результатів прогнозування та використання методу у підтримці прийняття рішень. Шостий етап методу завершує загальний цикл аналітичної обробки даних і спрямований на практичне використання отриманих результатів для підтримки управлінських рішень. Якщо попередні етапи забезпечували формування інформаційної основи, оцінювання значущості ознак, побудову інтегрального індексу та його прогнозування, то на цьому етапі здійснюється змістова інтерпретація отриманих прогнозних значень, просторове групування територій за рівнем ризику та визначення напрямів управлінського реагування.

Результатом роботи прогностичної моделі є множина прогнозованих значень інтегрального індексу ризику:

$$RI_i = f(X_i), \quad (19)$$

де RI_i — прогнозоване значення інтегрального індексу екологічного ризику для i -го муніципального об'єкта; $f(X_i)$ — функція прогнозованої моделі, що побудована на основі відібраних ознак.

Отримані значення інтерпретуються як кількісна характеристика екологічного навантаження від органічних відходів у межах конкретної території. Для практичного застосування результати групуються за рівнями ризику, що дозволяє сформувати просторову карту пріоритетності управлінських втручань.

Класифікація прогнозованих значень виконується за інтервалами, які подано у етапі 4 цього методу. Такий підхід дозволяє перейти від суто числових результатів до управлінсько зрозумілої класифікації територій. Для додаткового аналітичного контролю визначається середній прогнозований ризик по всій досліджуваній території:

$$\overline{RI} = \frac{1}{m} \sum_{i=1}^m RI_i, \quad (20)$$

де m — кількість муніципальних об'єктів.

Середнє значення дозволяє оцінити загальний рівень екологічної напруги в межах системи поводження з органічними відходами та використовується для порівняння різних територіальних груп.

Практичне значення цього етапу полягає в тому, що результати прогнозування стають основою для визначення управлінських пріоритетів. Території з високим прогнозованим ризиком потребують першочергового розвитку інфраструктури, оптимізації маршрутів збору органічних відходів, локального підсилення потужностей компостування або біогазового використання органічної фракції. Крім локальної інтерпретації, результати використовуються для цифрової підтримки рішень у муніципальному управлінні, де інтегральний індекс екологічного ризику є входним параметром для формування сценаріїв розвитку системи поводження з органічними відходами.

Для оцінки чутливості системи до змін окремих факторів використовується приріст прогнозованого індексу:

$$\Delta RI = RI_{new} - RI_{base}, \quad (21)$$

де RI_{new} — нове прогнозоване значення інтегрального індексу екологічного ризику після зміни параметрів; RI_{base} — базове значення інтегрального індексу екологічного ризику.

Це дозволяє оцінювати потенційний ефект управлінських дій ще до їхнього фактичного впровадження. Особливістю шостого етапу є інтеграція прогнозного результату в систему підтримки прийняття рішень, де модель перестає бути лише інструментом оцінювання і стає основою цифрового управління екологічними процесами. Таким чином, завершальний етап забезпечує практичну трансформацію математичного результату в прикладний інструмент муніципального управління, що є ключовою науковою новизною запропонованого методу.

Результати прогнозування інтегрального індексу екологічного ризику для підтримки управлінських рішень у системі муніципального поводження з органічними відходами

Вихідні дані для дослідження отримано від Львівського комунального підприємства «Зелене місто» (м. Львів), яке здійснює організацію збирання, транспортування та перероблення органічних відходів у межах територіальної громади. Отримані масиви охоплюють щоденні операційні показники, логістичні параметри, дані про навантаження на об'єкти обробки, а також екологічні вимірювання, зокрема концентрацію метану та допоміжні фізико-хімічні характеристики. Для прогнозування інтегрального індексу екологічного ризику *risk_index* на основі запропонованого методу було розроблено програмний модуль у середовищі Google Colab, який забезпечує повний цикл аналітичного опрацювання багатofакторного масиву даних: завантаження сформованої матриці спостережень, поділ вибірки на навчальну і тестову підмножини, навчання моделей, формування прогнозних значень та оцінювання точності результатів. Програмна реалізація виконана мовою Python із використанням бібліотек NumPy, pandas, Matplotlib та Scikit-learn, що дозволило сформувавши єдине обчислювальне середовище для порівняльного аналізу

моделей прогнозування. Для підвищення достовірності результатів застосовано п'ятикратну крос-валідацію, а також тестову вибірку у співвідношенні 80/20.

Початковий етап аналізу передбачав дослідження внутрішньої структури багатofакторного набору даних. Кореляційна матриця числових змінних дозволила виявити стійкі зв'язки між демографічними, просторовими, логістичними та екологічними параметрами (рис. 2).

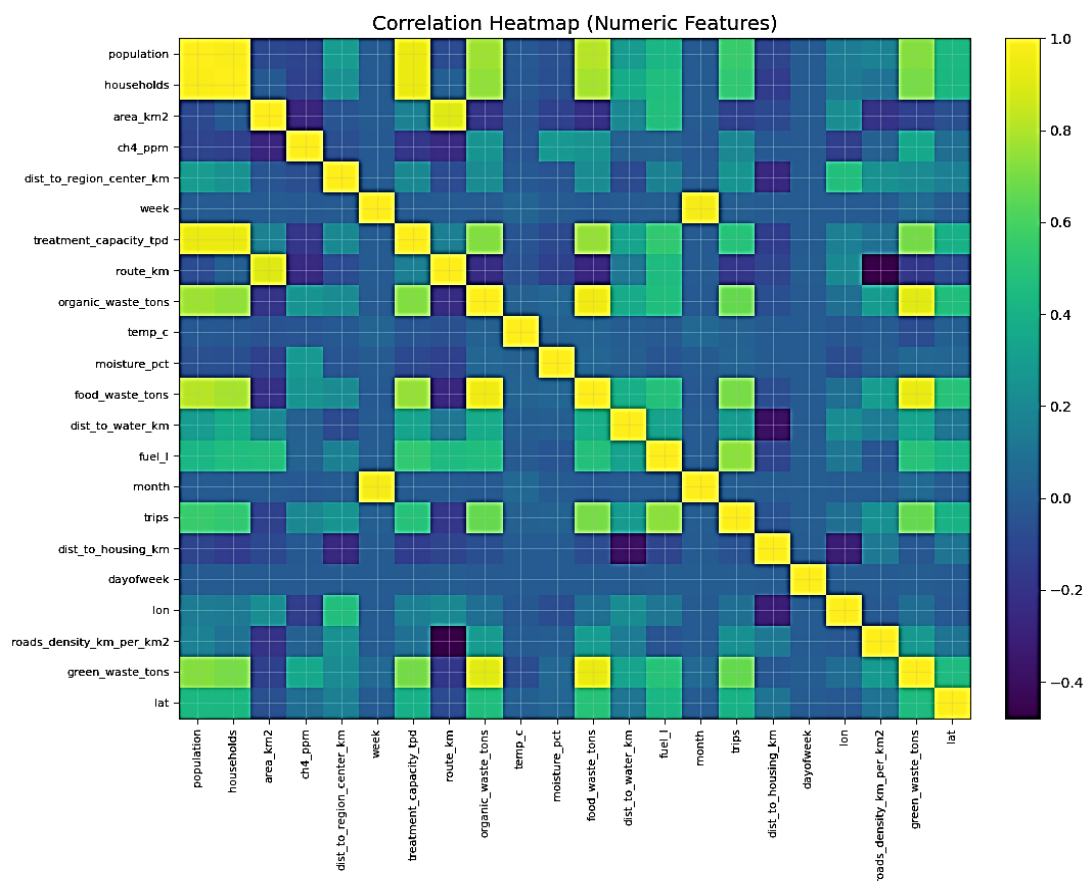


Рис. 2. Кореляційна матриця числових змінних багатofакторного набору даних

Установлено, що найбільш виражені позитивні залежності спостерігаються між показниками population, households, treatment_capacity_tpd та organic_waste_tons, що підтверджує їхню системоутворюючу роль у формуванні екологічного навантаження. Водночас частина просторових характеристик демонструє помірний або слабкий кореляційний вплив, що свідчить про необхідність додаткового інформаційного ранжування.

Для кількісного визначення інформаційної значущості ознак використано критерій взаємної інформації, який дозволив ранжувати змінні відповідно до їхнього внеску в пояснення інтегрального індексу ризику (рис. 3). Найвищі значення отримали tourism_index, dist_to_region_center_km, income_index, area_km2, dist_to_water_km, lon, urban_share, lat та population. Це означає, що просторово-економічні характеристики виявилися не менш значущими, ніж традиційні виробничі параметри системи поводження з органічними відходами.

Для систематизації інформації всі ознаки були об'єднані у функціональні блоки — демографічний, геопросторовий, логістичний, переробний, сенсорний, часовий і похідний (табл. 2). Такий підхід забезпечив структурну узгодженість подальшого прогнозування і дозволив формувати модель на основі логічно пов'язаних груп факторів.

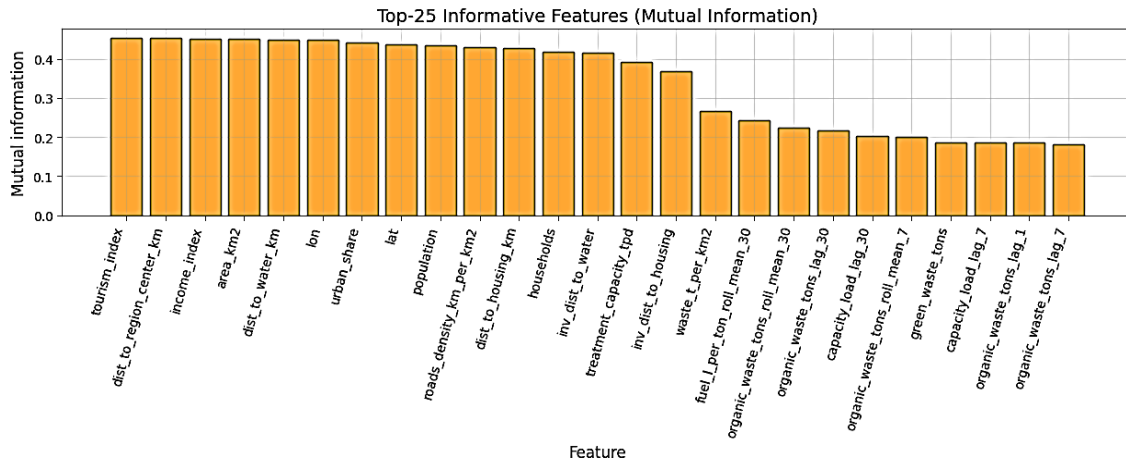


Рис. 3. Гістограма 25-ти найбільш інформативних ознак за критерієм взаємної інформації

Таблиця 2. Групування ознак багатфакторного набору даних за функціональними блоками

№	Група ознак	Приклади показників	Функціональне призначення в моделі
1	Демографічні	population, households	Відображають масштаб формування органічних відходів залежно від чисельності населення та кількості домогосподарств
2	Геопросторові	lat, lon, dist_to_water_km, urban_share	Характеризують просторове положення території, екологічну чутливість і рівень урбанізованості
3	Логістичні	route_km, trips, fuel_l	Описують параметри транспортного обслуговування системи збору органічних відходів
4	Показники перероблення	treatment_capacity_tpd, capacity_load	Відображають потужність і фактичне навантаження інфраструктури перероблення
5	Сенсорні показники	temp_c, moisture_pct, ph, ch4_ppm	Характеризують фізико-хімічний стан органічної маси та екологічне навантаження
6	Часові	year, month, week, dayofweek	Дозволяють враховувати сезонні та циклічні закономірності утворення відходів
7	Похідні та лагові	waste_kg_per_capita, lag/rolling features	Відображають похідні індикатори інтенсивності утворення відходів і часову інерційність процесів

Після формування ентропійно підготовленого набору ознак було реалізовано прогнозування інтегрального індексу ризику із використанням чотирьох моделей машинного навчання: Ridge Regression, Random Forest, Gradient Boosting та XGBoost (рис. 4). Порівняльний аналіз показав, що інтеграція ентропійного ранжування ознак суттєво підвищує пояснювальну здатність моделей.

Найкращий результат продемонструвала модель XGBoost, для якої коефіцієнт детермінації досяг значення — $R^2 = 0,940$. При цьому середньоквадратична похибка становила — $RMSE = 0,055$, а середня абсолютна похибка — $MAE = 0,039$.

Отримані результати свідчать, що ансамблеві алгоритми краще відтворюють нелінійну структуру взаємозв'язків між параметрами муніципальної системи поводження з органічними відходами.

Практичне використання прогнозованого індексу екологічного ризику *risk_index* полягає у формуванні цифрової аналітичної основи для пріоритизації територій, де необхідне інфраструктурне втручання, перегляд логістичних схем або локальна модернізація потужностей перероблення органічної фракції (табл. 3). У такому випадку прогнозний результат перетворюється на інструмент підтримки прийняття управлінських рішень, що орієнтований не лише на фіксацію поточного стану, а й на випереджувальне реагування.

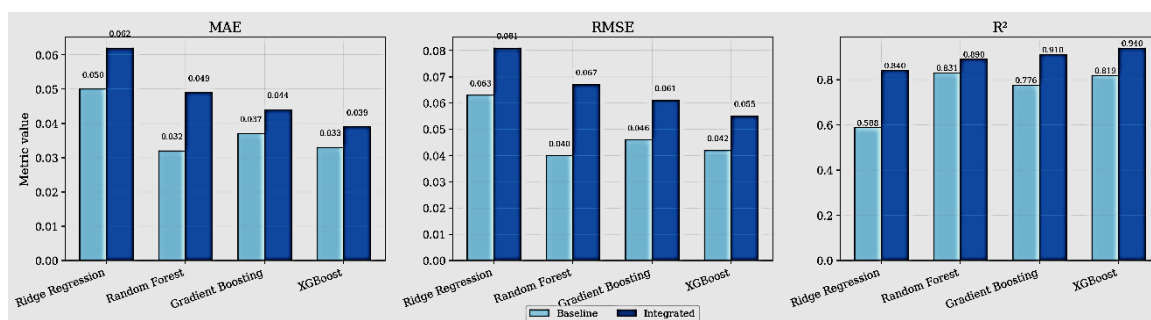


Рис. 4. Порівняльна ефективність моделей прогнозування інтегрального індексу екологічного ризику

Таблиця 3. Рекомендації щодо використання прогнозних результатів індексу екологічного ризику *risk_index* у муніципальному управлінні

Діапазон прогнозованого індексу <i>risk_index</i>	Рівень ризику	Управлінське рішення
0,00–0,33	низький	Поточний моніторинг і підтримання існуючої логістики
0,34–0,66	середній	Оптимізація маршрутів збору, коригування навантаження на інфраструктуру
0,67–1,00	високий	Першочергове інфраструктурне втручання, розширення потужностей перероблення

Таким чином, запропонований метод, який передбачає інтеграцію ентропійного ранжування ознак із використанням машинного навчання для створення моделі прогнозування дозволив сформувавши цілісний цифровий механізм оцінювання індексу екологічного ризику *risk_index*, який безпосередньо буде використано у системах муніципального управління органічними відходами та просторовому плануванні розвитку інфраструктури громад.

Висновки

Запропонований метод ентропійно-прогностичного аналізу інтегрального індексу екологічного ризику від муніципальних органічних відходів передбачає системне виконання 6 послідовних етапів, якими забезпечується формування багатфакторного набору даних, ентропійне оцінювання інформативності ознак, структу-

ризацію показників за функціональними групами, формування модельної матриці із підготовленими даними, машинне навчання для прогнозування екологічного ризику та інтерпретацію результатів для підтримки управлінських рішень. Новизна методу полягає в інтеграції ентропійного ранжування ознак із прогнозними алгоритмами машинного навчання для одночасного врахування інформаційної значущості ознак і нелінійних взаємозв'язків між просторовими, логістичними, екологічними та соціально-демографічними факторами, що впливають на ризик. Це дало можливість визначити найбільш інформативні предиктори, серед яких *tourism_index*, *dist_to_region_center_km*, *income_index*, *area_km²*, *dist_to_water_km* із значеннями взаємної інформації у межах віж 0,44 до 0,46, та сформуванню обґрунтовану основу для достатньо точного прогнозування інтегрального індексу ризику.

На основі даних Львівського комунального підприємства «Зелене місто» (м. Львів) виконано ентропійну підготовку багатofакторного масиву даних, у межах якої сформовано 7 функціональних груп ознак, визначено вагові коефіцієнти показників і реалізовано прогнозування інтегрального індексу ризику із використанням чотирьох моделей машинного навчання: Ridge Regression, Random Forest, Gradient Boosting та XGBoost. Встановлено, що ентропійна підготовка масиву даних перед навчанням моделей забезпечила підвищення пояснювальної здатності алгоритмів. Найкращий результат отримано для моделі XGBoost, для якої коефіцієнт детермінації зріс з 0,819 до 0,940, значення RMSE становило 0,055, а MAE — 0,039. Це дало можливість підвищити точність прогнозування інтегрального індексу ризику *risk_index* та забезпечити використання результатів для пріоритизації території підвищеного екологічного навантаження, оптимізації логістики збору органічних відходів і підтримки муніципальних управлінських рішень.

1. Xu P., Zheng H. A multi-AI approach to predicting municipal solid waste generation and recycling demand in Hong Kong. *Resources, Conservation and Recycling*. 2026. Vol. 225. Art. 108590. DOI:10.1016/j.resconrec.2025.108590.
2. Adu T.F., Mensah L.D., Rockson M.A.D., Kemausuor F. Forecasting municipal solid waste generation and composition using machine learning and GIS techniques: A case study of Cape Coast, Ghana. *Case Studies in Chemical and Environmental Engineering*. 2025. Art. 100218. DOI:10.1016/j.clwas.2025.100218.
3. Latif S.D., Hazrin N.A.B., Younes M.K., Ahmed A.N., Elshafie A. Evaluating different machine learning models for predicting municipal solid waste generation: A case study of Malaysia. *Environment, Development and Sustainability*. 2024. Vol. 26, № 5. P. 12489–12512.
4. Rao R., Kumar A., Singh P. AI-powered municipal solid waste management: A comprehensive review from generation to utilization. *Frontiers in Energy Research*. 2025. Vol. 13. Art. 1670679. DOI:10.3389/fenrg.2025.1670679.
5. Tryhuba I., Tryhuba A., Hutsol T., Cieszevska A., Andrushkiv O., Glowacki S., Brys A., Slobodian S., Tulej W., Sojak M. Prediction of biogas production volumes from household organic waste based on machine learning. *Energies*. 2024. Vol. 17, № 7. Art. 1786. DOI:10.3390/en17071786.
6. Tryhuba A., Tryhuba I., Bashynsky O., Hutsol T., Rozkosz A., Prokopova O. Forecasting of environmentally efficient agricultural production based on machine learning models. *Sustainability*. 2023. Vol. 15. Art. 11472. DOI:10.3390/su151511472.
7. Tryhuba A., Hutsol T., Glowacki S., Bashynsky O., Rozkosz A., Prokopova O. Intelligent forecasting of resource consumption in agricultural production projects under uncertainty. *Applied Sciences*. 2024. Vol. 14. Art. 3128. DOI:10.3390/app14073128.

8. Chang T.-W., Huang Y.-P., Sandnes F.E. Efficient entropy-based features selection for image retrieval. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. 2009. DOI:10.1109/ICSMC.2009.5346032.
9. Nourani V., Baghanam A.H., Samadi E., Uzelaltinbulat S. Predicting municipal solid waste generation using artificial intelligence: A hybrid approach of entropy analysis and SHAP for optimal feature selection. *Waste Management*. 2025. Vol. 205. Art. 115012. DOI:10.1016/j.wasman.2025.115012.
10. Лапчук А.С., Пригун О.В., Манько Д.Ю., Горбов І.В., Бородин Ю.О. Умови ефективного застосування пасивних методів зменшення спеклів і метод декогерентного фокусування в лазерних освітлювальних системах і проекторах. *Ресстрація, зберігання і оброб. даних*. 2025. Т. 27, № 2. С. 3–26. DOI:10.35681/1560-9189.2025.27.2.345498.
11. Lee J.-S., Shin D.-C. Prediction of waste generation using machine learning: A regional study in Korea. *Urban Science*. 2025. Vol. 9, № 8. Art. 297. DOI:10.3390/urbansci9080297.
12. Alkhayrat A., Aljnidi A., Aljoumaa A. A comparative dimensionality reduction study in telecom customer segmentation using deep learning and explainable artificial intelligence. *Journal of Big Data*. 2024. Vol. 11. Art. 44. DOI:10.1186/s40537-020-0286-0.
13. Tryhuba A., Koval N., Tryhuba I., Boiarchuk O. Application of SARIMA models in information systems forecasting seasonal volumes of food raw materials of procurement on the territory of communities. In: Proceedings of the International Conference on Information Technologies in Project Management. P. 64–75. URL: <https://ceur-ws.org/Vol-3295/paper6.pdf>
14. Suknark P., Youwai S., Kitkobsin T., Towprayoon S. Explainable artificial intelligence model for evaluating shear strength parameters of municipal solid waste across diverse compositional profiles. arXiv preprint. 2025. DOI:10.48550/arXiv.2502.15827.
15. Fung P.L., Savadkoochi M., Zaidan M.A., Niemi J.V., Timonen H., Pandolfi M., Alastuey A., Querol X., Hussein T., Petäjä T. Constructing transferable and interpretable machine learning models for black carbon concentrations. *Environment International*. 2024. Vol. 184. Art. 108449. DOI:10.1016/j.envint.2024.108449.
16. Tryhuba A., Bashynskiy O., Medvediev Y., Slobodian S., Skorobogatov D. Justification of models of changing project environment for harvesting grain, oilseed and legume crops. *Independent Journal of Management & Production*. 2019. Vol. 10, № 7. P. 658–672. DOI:10.14807/ijmp.v10i7.922.
17. Lips S., DeYoung S., Schönsteiner M., Lens H. Closed-loop identification of a MSW grate incinerator using Bayesian optimization for selecting model inputs and structure. *Control Engineering Practice*. 2024. Vol. 153. Art. 106075. DOI:10.1016/j.conengprac.2024.106075.
18. El Jaouhari A., Samadhiya A., Kumar A., Mulat-Weldemeskel E., Luthra S., Kumar R. Turning trash into treasure: Exploring the potential of AI in municipal waste management – an in-depth review and future prospects. *Journal of Environmental Management*. 2024. DOI:10.1016/j.jenvman.2024.123658.
19. Wang Y., Chen X., Li J. Urban municipal organic waste forecasting using machine learning and environmental indicators. *Waste Management*. 2024. Vol. 182. P. 95–108. DOI:10.1016/j.wasman.2024.01.018.
20. Malanchuk O., Tryhuba A., Tryhuba I., Bandura I. A conceptual model of adaptive value management of project portfolios of creation of hospital districts in Ukraine. In: Proceedings of the International Conference on Information Technologies in Project Management. 2020. P. 82–95. URL: <https://ceur-ws.org/Vol-3453/paper8.pdf>
21. Khan M., Alhussain A., Alotaibi S. Municipal solid waste generation forecasting using deep learning and socio-economic indicators. *Journal of Cleaner Production*. 2023. Vol. 430. Art. 139684. DOI:10.1016/j.jclepro.2023.139684.
22. Lee J.S. Prediction of waste generation using machine learning. *Urban Science*. 2025. Vol. 9, № 8. Art. 297. DOI:10.3390/urbansci9080297.
23. Watkins J., Bertagna L., Harper G. Entropy-based feature selection for capturing impacts in Earth system models with abrupt forcing. *Journal of Computational and Applied Mathematics*. 2026. Vol. 471. Art. 116724. DOI:10.1016/j.cam.2025.116724.
24. Fung P.L., Savadkoochi M., Zaidan M.A. Constructing transferable and interpretable machine learning models for black carbon concentrations. *Environment International*. 2024. Vol. 184. Art. 108449. DOI:10.1016/j.envint.2024.108449.

25. Nourani V., Baghanam A.H., Samadi E., Uzelaltinbulat S. Predicting municipal solid waste generation using artificial intelligence: A hybrid approach of entropy analysis and SHAP for optimal feature selection. *Waste Management*. 2025. Vol. 205. Art. 115012. DOI:10.1016/j.wasman.2025.115012.
26. Tryhuba I., Tryhuba A., Hutsol T., Cieszewska A., Andrushkiv O., Glowacki S., Bryś A., Slobodian S., Tulej W., Sojak M. Prediction of biogas production volumes from household organic waste based on machine learning. *Energies*. 2024. Vol. 17. Art. 1786. DOI:10.3390/en17071786.
27. Tryhuba A., Hutsol T., Česna J., Glowacki S., Bashynsky O., Tulej W., Sojak M. Optimizing energy systems of livestock farms with computational intelligence for achieving energy autonomy. *Scientific Reports*. 2025. Vol. 15. Art. 10777. DOI:10.1038/s41598-025-92836-6.
28. Каденко С.В., Циганок В.В., Андрійчук О.В., Карабчук А.В. Аналіз інструментарію підтримки прийняття рішень у контексті вирішення задач стратегічного планування. *Реєстрація, зберігання і оброб. даних*. 2020. Т. 22, № 2. С. 77–91. DOI:10.35681/1560-9189.2020.22.2.211281.
29. Циганок В.В. Удосконалення методу цільового динамічного оцінювання альтернатив та особливості його застосування. *Реєстрація, зберігання і оброб. даних*. 2013. Т. 15, № 1. С. 90–99. DOI:10.35681/1560-9189.2013.15.1.103369.
30. Бойко О. Сучасні методи штучного інтелекту для виявлення пропаганди в тексті. *Реєстрація, зберігання і оброб. даних*. 2025. Т. 27, № 1. С. 120–131. DOI:10.35681/1560-9189.2025.27.1.336147.
31. Kavrin D., Subbotin S. Bagging-based instance selection for instance-based classification. In: *Proceedings of the Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020)*. CEUR Workshop Proceedings. 2020. Vol. 2608. P. 703–714. URL: <https://ceur-ws.org/Vol-2608/paper58.pdf>
32. Leoshchenko S., Subbotin S., Borovikov A., Gofman Y. Transfer training tools and methods for diagnostic tasks. In: *Proceedings of the Eighth International Workshop on Computer Modeling and Intelligent Systems (CMIS-2025)*. CEUR Workshop Proceedings. 2025. Vol. 3988. P. 312–324. URL: <https://ceur-ws.org/Vol-3988/paper28.pdf>
33. Kondysiuk O., Bashynsky O., Dembitskyi V., Myskovets I. Formation and risk assessment of stakeholders value of motor transport enterprises development projects. In: *Proceedings of the International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*. 2021. P. 303–306. DOI:10.1109/CSIT52700.2021.9648739.
34. Padyuka R., Tymochko V., Lub P. Mathematical model for forecasting product losses in crop production projects. In: *CEUR Workshop Proceedings*. 2022. Vol. 3109. P. 25–31.
35. Zachko O., Grabovets V., Pavlova I., Rudynets M. Examining the effect of production conditions at territorial logistic systems of milk harvesting on the parameters of a fleet of specialized road tanks. *Eastern-European Journal of Enterprise Technologies*. 2018. Vol. 5. P. 59–69. DOI:10.15587/1729-4061.2018.142227.

Надійшла до редакції 15.02.2026