

УДК 519.7+004.8

DOI [https://doi.org/10.24144/2616-7700.2024.44\(1\).93-105](https://doi.org/10.24144/2616-7700.2024.44(1).93-105)В. Г. Гнатієнко<sup>1</sup>, Г. М. Гнатієнко<sup>2</sup>, О. Л. Зозуля<sup>3</sup>, В. Є. Снитюк<sup>4</sup>

<sup>1</sup> Київський національний університет ім. Т. Шевченка,  
магістрант кафедри інтелектуальних технологій  
vladgnat1483@gmail.com  
ORCID: <https://orcid.org/0009-0000-2678-5158>

<sup>2</sup> Київський національний університет ім. Т. Шевченка,  
заступник декана факультету інформаційних технологій з наукової роботи,  
кандидат технічних наук  
g.gna5@ukr.net  
ORCID: <https://orcid.org/0000-0002-0465-5018>

<sup>3</sup> ТОВ «Сингента»,  
керівник групи технічних радників,  
кандидат біологічних наук  
alexandr.zozulya@syngenta.com  
ORCID: <https://orcid.org/0000-0003-3500-3423>

<sup>4</sup> Київський національний університет ім. Т. Шевченка,  
декан факультету інформаційних технологій,  
доктор технічних наук, професор  
snytyuk@knu.ua  
ORCID: <https://orcid.org/0000-0002-9954-8767>

## МЕТОД ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР З ВИКОРИСТАННЯМ МУЛЬТИФАКТОРНОГО АНАЛІЗУ ТА НЕЙРОННИХ МЕРЕЖ

Дослідження присвячене розробці інтелектуальної технології прогнозування врожайності з використанням супутникових, геоінформаційних даних та кліматичних показників. Використання сучасних методів машинного навчання надає представникам аграрного сектору стратегічні переваги, дозволяючи знижувати ризики від надмірного використання пестицидів та сприяти сталому розвитку сільського господарства. У статті пропонується здійснювати оптимізацію процесу переджнивного підсушування рослин з метою прискорення достигання і полегшення збирання врожаю (десикацію) соняшнику, використовуючи моделювання обсягів можливих втрат врожаю на основі даних, одержаних в період вегетації культури. Така пропозиція є актуальною, оскільки її використання забезпечує підвищення точності прогнозів та ефективність управлінських рішень разом із зменшенням витрат на обробку полів.

**Ключові слова:** сільськогосподарські культури, десикація, прогнозування, супутникові дані, кліматичні показники, вегетаційні індекси, нейронні мережі.

**1. Вступ.** Активний розвиток цифрової агрономії відкриває широкі перспективи для інтенсифікації розвитку аграрного сектору, одночасно породжуючи низку складних задач і викликів. На тлі змін клімату, коливань ринкових цін та зростаючих вимог до ефективності використання природних ресурсів, стандартизація вимог законодавства України щодо захисту рослин до Європейських норм у плануванні бюджету та оптимізація процесів догляду за рослинами стає все більш актуальною. Сучасний розвиток цифрових технологій, штучних нейронних мереж, штучного інтелекту, нових підходів до статистичної обробки даних є передумовою переходу виробників на новий рівень ведення сільськогосподарського виробництва.

Недосконалість та неможливість застосування традиційних статистичних підходів, які надають лише приблизну оцінку врожайності, стають особливо помітними на фоні потреб сучасного аграрного виробництва. Хоча ці методи забезпечують аналіз поточного стану та прогнозування майбутньої врожайності, вони не дозволяють задовольнити потребу в точному та деталізованому плануванні. У той же час, штучний інтелект з його можливостями глибокого аналізу даних і машинного навчання, відкриває нові горизонти для значного підвищення точності прогнозування врожайності. Відповідні технології мають великий потенціал для революціонізації агрономії, пропонуючи інноваційні рішення для оптимізації агротехнічних заходів та підвищення ефективності виробництва. Перехід до використання цих передових технологій вимагає не лише розвитку нового інструментарію та методик, але й глибокого переосмислення підходів до управління аграрними процесами.

**2. Аналіз останніх досліджень і публікацій.** Сучасні системи прогнозу врожайності з використанням штучного інтелекту охоплюють широкий спектр методів і технологій. Наведемо кілька важливих напрямків досліджень у цій галузі.

1. Використання глибоких нейронних мереж дозволяє досить точно прогнозувати врожайність, аналізуючи дані про погодні умови і характеристики ґрунту та демонструючи середню точність 85–89%. Однак, головним недоліком такого підходу є його обмеження щодо прогнозування на рівні всього поля, що не дозволяє враховувати локальні мікрокліматичні та ґрунтові варіації, що є важливими для деталізованого прогнозування [1].
2. Машинне навчання з використанням традиційних алгоритмів забезпечує високу точність прогнозування загальної врожайності поля. Однак, попри теоретичні можливості деталізованого аналізу, в рамках цього дослідження детальний прогноз по окремих ділянках не описується та не реалізується [2].
3. Дослідження, в яких застосовуються рекурентні нейромережі з використанням навчання з підкріпленням для прогнозування врожайності, демонструють середню точність на рівні 93.7%. Цей метод фокусовано на визначенні загальної врожайності всього поля, але здійснити прогнозування за окремими ділянками не можливо [3].
4. У дослідженнях, де використовувалась стратифікована випадкова вибірка для збору урожаю картоплі та емпіричні рівняння на основі індексів NDVI та SAVI, досягнуто значних результатів з похибками 3.8%–7.5%, однак тестова вибірка належала тому ж полю, що і тренувальна [4].

Сучасні підходи до прогнозування врожайності, що використовують технології штучного інтелекту, безсумнівно, є підставою для досягнення значних результатів при обробці великих обсягів даних та наданні достовірних прогнозів на рівні цілого поля. Усі методи, які ґрунтуються на аналізі складної взаємодії між генетичними характеристиками рослин, погодними умовами, властивостями ґрунту та іншими зовнішніми факторами, демонструють високу точність прогнозування.

Попри свою ефективність у визначенні загальних трендів врожайності, підходи, які розглянуто вище, мають істотні обмеження, особливо коли мова йде про деталізацію до рівня окремих ділянок поля, а не всього поля в цілому.

Основною проблемою є те, що більшість існуючих моделей орієнтовані на прогнозування сумарної врожайності всього поля, не враховуючи внутрішньопольові варіації, які можуть мати критичне значення для точного управління агротехнічними заходами. Це обмеження робить неможливим формування детальної карти врожайності, що, в свою чергу, обмежує потенціал застосування таких систем для розв'язання ряду ключових задач. Зокрема, оптимізація диференційованого внесення добрив та засобів захисту рослин, підтримки водного режиму, проведення біологічних досліджень для вивчення впливу різних комбінацій агротехнічних факторів на продуктивність конкретних ділянок ґрунту, залишається поза межами можливостей цих технологій. У дослідженні, де було сформовано емпіричні рівняння для прогнозування [4], хоч і було реалізовано метод деталізованого прогнозування, тестові вибірки належали тим же самим полям, на яких було сформовано тренувальні вибірки, що значно обмежує потенційні можливості застосування моделі через її недостатню універсальність. Розвиток методів, здатних забезпечити деталізований аналіз на рівні локальних ділянок, стає пріоритетним напрямком удосконалення систем прогнозування врожайності, відкриваючи нові перспективи для точного землеробства.

**3. Постановка задачі.** Основна мета цього дослідження полягає у розробці технології прогнозування врожайності та визначення можливих втрат внаслідок нерівномірного дозрівання. Ці втрати можуть бути суттєво зменшені шляхом десикації посівів. У кожному конкретному випадку виникає питання доцільності цього додаткового агротехнічного заходу. Його вирішення пов'язане із прогнозуванням врожайності на окремих ділянках поля, прогнозуванням втрат на кожній ділянці, оптимізації умов посіву та догляду за рослинами, включаючи густоту посіву, час та інтенсивність внесення гербіцидів та фунгіцидів.

Для побудови моделі визначено предиктори, які можуть впливати на нерівномірність дозрівання, а саме — дата посіву, погодні умови, вологість, ФАО (або групу стиглості гібриду), попередники тощо. Такий прогноз дає можливість фермеру оцінити економічну доцільність агротехнічного захисту, який планується, застосувати вибіркочну обробку окремих ділянок, зменшуючи пестицидне навантаження на навколишнє середовище.

Результатом дослідження має стати система прогнозування врожайності з деталізацією до локальних ділянок, яка інтегрує супутникові знімки, метеорологічні дані, інформацію про рослину та її догляд (гібрид, густоту посіву, типи хімічної обробки, декілька видів вегетаційних індексів — GLI, CLg, CLr, NDWI, NDVI тощо).

Для побудови деталізованої карти врожайності виконується розбиття поля на окремі ділянки, на кожній з яких виконується окремий прогноз. Позначимо реальне значення врожайності  $i$ -ої ділянки поля  $Y$  як  $y_i \in Y$ ,  $i = 1, \dots, n$ , а прогнозоване —  $\hat{y}_i = f_0(x_i)$ ,  $i = 1, \dots, n$ , де  $x_i \in X$ , — вектор вхідних даних, що описує стан  $i$ -ої ділянки  $i = 1, \dots, n$ , впродовж періоду дозрівання,  $X$  — матриця даних для всіх ділянок поля  $Y$ ;  $\theta$  — параметри моделі прогнозування врожайності  $Y$ ;

$\theta$  — параметри моделі прогнозування;

$f$  — функціональна залежність між вхідними даними стану поля та врожайністю, що встановлюється методами машинного навчання.

Потрібно ідентифікувати залежність, яка є розв'язком задачі

$$L(f_{\theta}(x_i), y_i) \rightarrow \min, \quad i = 1, \dots, n,$$

де  $n$  — кількість ділянок на полі  $Y$ ,  $L$  — функція втрат, що відображає відхилення прогнозованих значень від реальних, та здійснити прогнозування врожайності при заданих вхідних даних.

**Математична модель** Задача прогнозування врожайності є надзвичайно складною, порівнюваною із прогнозуванням погоди, і вимагає не тільки врахування великої кількості параметрів, але й ідентифікації ключових факторів, від яких найбільше залежить результат.

На врожайність впливає велика кількість параметрів: хімічний склад та структура ґрунту, його вологість, рівень рН, внесені типи добрив, особливості їх застосування тощо. Важливими факторами є погодні умови: температура повітря, рівень опадів, вологість, інтенсивність сонячного світла, а також присутність та активність шкідників і хвороб. Не менш важливими є агротехнічні заходи: обробка ґрунту, дотримання сівозмін, методи посіву та методи збирання врожаю. Крім того, необхідно враховувати генетичні особливості посівного матеріалу, його стійкість до хвороб та адаптивність до погодних умов.

Нагадаємо, що основною задачею даного дослідження є прогнозування врожайності кожної ділянки поля на основі аналізу детальних даних про стан рослин, що включають дані, отримані з супутникових знімків у різних спектрах, які конвертуються у вегетаційні індекси NDVI, NDWI, CLg, CLr, GLI [5]. Важливими є також метеорологічні дані: температура, кількість опадів, швидкість та напрямки вітру, хмарність, сонячна радіація та атмосферний тиск. Доповнюються ці дані інформацією про агротехнічні заходи, включаючи обробку поля гербіцидами та фунгіцидами, тип гібрида соняшника (раннього чи пізнього) та густоту посіву. Для навчання та валідації моделі використовується тренувальний датасет, в якому для кожної ділянки поля надано дані про урожайність в тонах на гектар.

Введемо умовні позначення для подальшого опису методу.

Нехай  $X_i$  — матриця даних для  $i$ -го поля, елементи матриці  $x_{ijkl} \in X_i$ ,  $n$  — кількість полів у навчальній вибірці; дні проведення спостережень  $j \in J = \{j_1, \dots, j_m\}$ ;  $m$  — кількість днів спостережень, що були проведені для поля;

Кожній ділянці відповідає рядок матриці

$$l \in L = \{NDVI, NDWI, GLI, CLr, CLg, \text{wind speed}, \text{seeding density}, \dots\}$$

— вхідні дані.

Множина вхідних даних має значний обсяг. Структуру даних наведено на рисунку 1. Дані за усі дні спостережень складають разом багатовимірний масив розмірності  $n \times |L| \times m$ .

Щоб зменшити розмірність вхідного вектора даних та оптимізувати процес прогнозування, проводиться попередній аналіз даних та відбір найбільш значущих ознак (feature selection). Одним з ключових інструментів у цьому процесі є кореляційний аналіз, який допомагає виявити статистичні зв'язки між різними ознаками. З допомогою кореляційного аналізу виявляють ознаки, які сильно

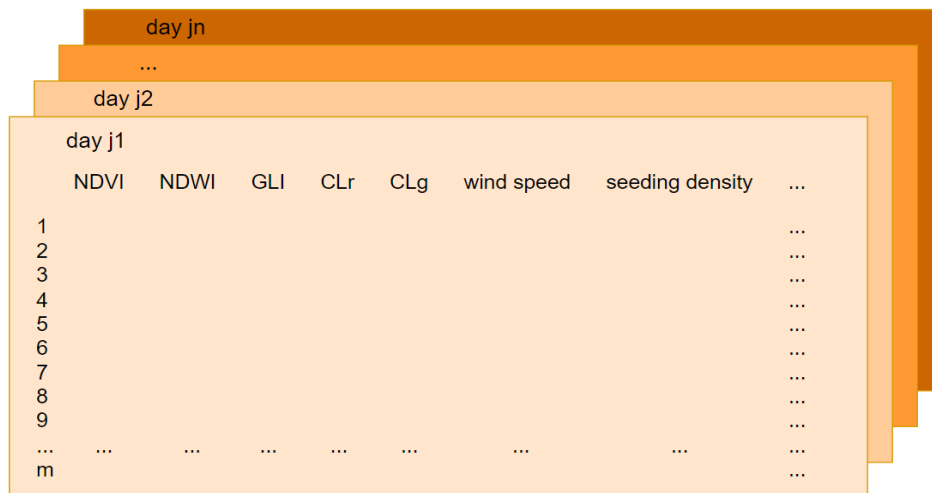


Рис. 1. Схематичне представлення структури навчального набору даних до проведення попередньої обробки.

корелюють між собою. Серед таких ознак експертним шляхом з допомогою експертів з аграрного сектору визначається тільки одна ознака, відкидаючи інші для зменшення розмірності датасету. Це дозволяє підвищити ефективність моделі, зберігаючи тільки найінформативніші ознаки.

На етапі попередньої обробки виконуються наступні кроки.

Крок 1. Видалення викидів для кожного дня окремо для кожного поля із застосуванням методу z-score [6, 7].

Значення z-score для кожного елемента визначається за формулою:

$$z_{ijkl} = \frac{x_{ijkl} - \mu_{ijl}}{\sigma_{ijl}},$$

де  $x_{ijkl} \in X_i$ ;  $\mu_{ijl} = \sum_{k=1}^m x_{ijkl}/m$  — середнє значення  $l$ -го параметра за день  $j$ -ий в матриці  $X_i$ ;  $\sigma_{ijl} = \sqrt{\sum_{k=1}^m (x_{ijkl} - \mu_{ijl})^2/m}$  — стандартне відхилення  $l$ -го параметра за  $j$ -ий день в матриці  $X_i$ ,  $z_{ijkl}$  — z-оцінка елемента  $x_{ijkl}$ .

**Евристика Е1.** Будемо вважати елемент  $x_{ijkl}$  викидом, якщо значення  $|z_{ijkl}| > 2.9$ . У такому випадку весь  $k$ -ий рядок за  $j$ -ий день матриці  $X_i$  видаляється з навчального набору.

Крок 2. Для кожної матриці  $X_i, i = \overline{1, g}$ , для кожної послідовності елементів  $x_{ijkl} \in X_i, j \in J$  обчислюються агреговані дані за формулами:

$$x_{ikl}^{\min} = \min_{j \in J} (x_{ijkl}), \quad x_{ikl}^{\text{mean}} = \frac{1}{|J|} \sum_{j \in J} x_{ijkl}, \quad x_{ikl}^{\max} = \max_{j \in J} (x_{ijkl}),$$

утворюючи матрицю агрегованих значень  $X'_i$ .

Крок 3. Об'єднання даних у загальний датасет  $X = \bigcup_{i=\overline{1, g}} X'_i$ .

Крок 4. Повторне видалення викидів на об'єднаному датасеті  $X$ :

$$z'_{ql} = \frac{x'_{ql} - \mu'_l}{\sigma'_l},$$

де  $x'_{ql} \in X$  — елемент матриці, що відповідає рядку  $q$  і параметру  $l$ ;

$\mu'_l = \sum_{q=1}^s x'_{ql}/s$  — середнє значення параметра  $l$  в матриці  $X$ ;

$\sigma'_l = \sqrt{\sum_{q=1}^s (x'_{ql} - \mu'_l)^2/s}$  — стандартне відхилення параметра  $l$  в матриці  $X$ ;

$z'_{ql}$  —  $z$ -оцінка елемента  $x'_{ql}$ .

**Евристика Е2.** Будемо вважати елемент  $x'_{ql}$  викидом, якщо  $|z'_{ql}| > 2.9$ . У такому випадку весь рядок  $q$  матриці  $X$  видаляється з навчального набору.

Повторне видалення викидів є важливим кроком, оскільки видалення викидів окремо в межах кожного поля не гарантує відсутності неправильних даних у зібраному наборі. При об'єднанні різних полів можливі ситуації, коли дані, які вважалися нормальними для одного поля, стають аномальними у контексті загального набору через різницю в масштабах, розподілах чи інших характеристиках. Тому необхідно знову видаляти викиди, щоб забезпечити консистентність та однорідність всіх зібраних даних.

У дослідженні для категоріальних ознак, таких як типи гібридів та рослини-попередники, було застосовано метод one-hot encoding. Цей метод дозволяє перетворити категоріальні дані у форму, яка придатна для обробки алгоритмами машинного навчання. Для кожної категоріальної ознаки створюється набір нових змінних (стовпців), де кожна змінна представляє одну категорію. Якщо в оригінальному записі спостерігалася певна категорія, то відповідний стовпчик для цієї категорії отримує значення 1, а всі інші — 0.

Для внесення даних про дати гербіцидних та фунгіцидних обробок, які є частиною застосування захисних засобів у рослинництві, було реалізовано метод, схожий на one-hot encoding. Замість використання безпосередньо кількості днів від дати посіву до моменту обробки, дані були представлені через бінарні стовпчики, що відображають певні діапазони днів проведення спостережень.

**Евристика Е3.** Експертами було встановлено типові часові діапазони для внесення фунгіцидів (34–82 дні від дати посіву) та гербіцидів (24–62 дні).

**Евристика Е4.** Ці діапазони днів проведення спостережень було поділено на окремі категорії. Для фунгіцидів це стовпчики  $\{34, 46, 58, 70, 82\}$ , а для гербіцидів —  $\{24, 33, 43, 52, 62\}$ .

**Евристика Е5.** Кожна дата внесення хімікатів віднесена до найближчої категорії. Наприклад, якщо фунгіциди були застосовані на 48-й день після посіву, то вектор для цього випадку буде складатися з елементів  $\{0, 1, 0, 0, 0\}$ , що відповідає категорії 46.

У результаті застосування описаних методів обробки даних для кожної ділянки поля було створено вектор інформації, який описує її розвиток впродовж періоду дозрівання. Для навчання прогнозу моделі була використана Light Gradient Boosting Machine (LightGBM) [8], яка є ансамблевою моделлю, що послідовно будує та ускладнює дерева рішень, використовуючи градієнтний спуск для оптимізації втрат. Використання LightGBM у цьому дослідженні дозволи-

ло ефективно здійснювати прогнозування врожайності, надаючи можливість детально аналізувати кожен окрему ділянку поля ізольовано.

Розглядаючи аспекти аналізу інформації, крім традиційного векторного представлення моніторингових значень, можна виокремити візуальний аспект, який дає змогу оцінювати великі території на основі образів. Це доповнює числові дані, надаючи ширшу картину стану сільськогосподарських угідь [9, 10]. Об'єднання цих підходів сприяє створенню синергетичного ефекту, розширює уявлення про об'єкт дослідження та збільшує точність прогнозів. Для забезпечення контекстної інформації було розроблено спеціалізовану модель комп'ютерного зору на основі архітектури U-net, яка аналізує залежності на великих територіях, ідентифікуючи високопродуктивні зони та області з потенційно нижчою врожайністю.

У процесі роботи моделі комп'ютерного зору вхідною інформацією для обробки є прямокутні зображення (патчі), отримані в результаті розбиття великих полів на менші сегменти.

Оскільки сільськогосподарські поля мають різноманітні розміри, у цьому дослідженні для ефективної обробки даних було застосовано метод розбиття зображень на менші частини, відомі як патчі. Такий підхід дозволяє забезпечити деталізовану сегментацію кожної частини поля окремо, після чого отримані патчі зіставляються разом, утворюючи цілісне сегментоване зображення поля. Цей процес не лише спрощує обробку великих ділянок землі, але й підвищує точність сегментації за рахунок детального аналізу кожного фрагменту.

У процесі накладання патчів було використано метод з ваговими коефіцієнтами, який дозволяє більш гладко інтегрувати частини зображення між собою. При цьому кожен піксель на ділянках, які перекриваються, отримує вагу залежно від його відстані до центру патча. Ця процедура сприяє м'якшому і більш природньому переходу між сегментами. Такий підхід мінімізує ризик різких відмінностей на границях патчів, забезпечуючи високу якість кінцевого сегментованого зображення і точне відображення варіативності врожайності на полях. Формула для визначення вагового коефіцієнта  $w(x, y)$  для пікселя, розташованого на відстані  $(x, y)$  від центру патча, може бути представлена таким чином:

$$w(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}},$$

де  $x$  та  $y$  — відстані від пікселя до центру патча за відповідними осями;  $\sigma$  — параметр, що визначає ширину "дзвона" гаусівської кривої, тобто рівень згладжування на краях патчів.

Враховуючи, що краї полів зазвичай не є прямими, і кожне поле характеризується своєю унікальною формою, виникає необхідність ефективної обробки патчів, що частково охоплюють межі поля і виходять за них. Для вирішення цієї проблеми і оптимізації процесу сегментації були впроваджені маски, які накладаються на патчі.

Ці маски виконують функцію індикаторів, вказуючи моделі, які пікселі в патчі належать до поля, а які знаходяться за його межами і не повинні враховуватися при аналізі. Таким чином, маски дозволяють моделі фокусуватися лише на релевантних ділянках зображення, ігноруючи ті частини патчів, що виходять за межі поля. Це значно підвищує точність і ефективність сегментації,

оскільки зменшується кількість шуму та нерелевантних даних, які обробляються моделлю.

Модель U-net, розроблена для виконання сегментації земельних ділянок, ефективно аналізує умови розвитку рослин на великих масштабах, враховуючи комплексні змінні та індивідуальні особливості кожної ділянки (рис. 2 [7]). Після завершення сегментації, кожна ідентифікована ділянка поля проходить через додатковий етап аналізу за допомогою алгоритму LightGBM, який виконує прогнозування з урахуванням конкретного сегменту, до якого належить ділянка.

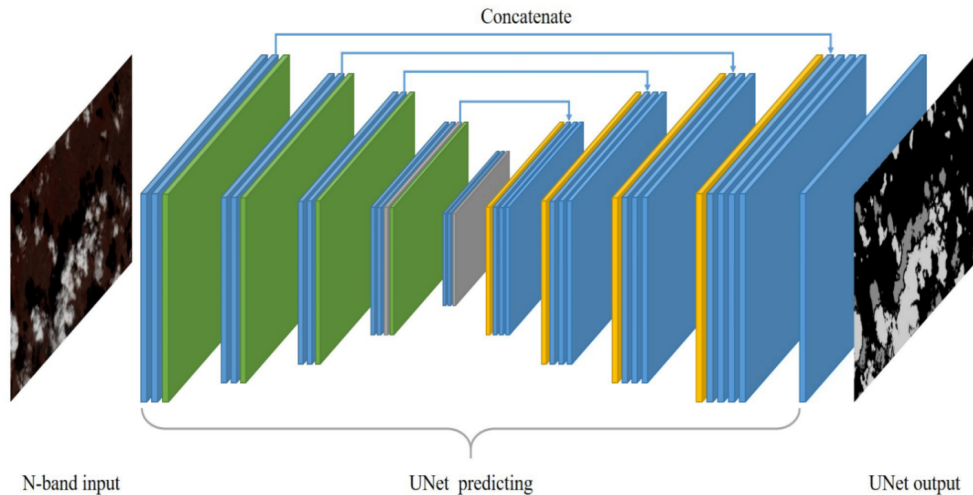


Рис. 2. Приклад виконання сегментації на основі супутникового зображення [7].

**4. Одержані результати дослідження.** У межах експериментальних досліджень компанії «Сингента» для перевірки точності прогнозування врожайності було виділено декілька полів, які не було включено до навчальної вибірки. Кожне поле розділене на окремі ділянки, для яких модель виконує прогнозування врожайності індивідуально. Прогноз для кожної ділянки порівнюється з фактичними даними збору врожаю, на основі чого обчислюється корінь середньоквадратичної помилки (RMSE), що дозволяє оцінити точність моделі.

Далі, врожайність кожної ділянки у тонах на гектар переводиться в тони, шляхом множення на площу однієї ділянки. Таким чином, обчислюється сумарна врожайність всього поля за прогнозованими даними та порівнюється з фактичною сумарною врожайністю. Точність сумарного прогнозу врожайності поля оцінюється шляхом розрахунку відсоткового відхилення прогнозованої врожайності від реальної.

Хоча ключовим показником є точність сумарної врожайності, точність прогнозування для окремих ділянок також має велике значення для подальших досліджень. Завдяки цьому можна масштабувати та універсалізувати систему, забезпечивши високу точність як у загальному прогнозі, так і в індивідуальному застосуванні до окремих ділянок. Це дозволить вирішувати широке коло завдань, таких як оптимізація диференційованого внесення препаратів, оцінка



впливу точкової обробки поля різними агрохімічними речовинами, оцінка локального впливу різних комбінацій факторів, включно з рельєфом, відбитим сонячним випромінюванням та станом ґрунту кожної конкретної ділянки.

У рамках дослідження ефективності системи прогнозування врожайності було отримано результати, які представлено в таблиці 1.

Таблиця 1.

Результати дослідження ефективності системи прогнозування врожайності

Поле	MAE	Точність	Прогнозований врожай (т)	Фактичний врожай (т)	Площа (га)
Flora_Baba_22	0.360007	98.299489	189.184446	192.400052	101
East-West_Serby_26_23	0.722826	96.284523	89.576487	87.459699	26.6
East-West_Serby_37_23	0.585210	92.651682	98.978797	106.251481	37
East-West_Serby_57_23	0.705239	93.220749	209.482780	195.217092	57.4
East-West_Serby_69_23	0.724573	95.723041	253.647770	242.73062	69
Zhuravske_Field_2_22	0.548212	86.796210	126.477451	109.757490	29.9

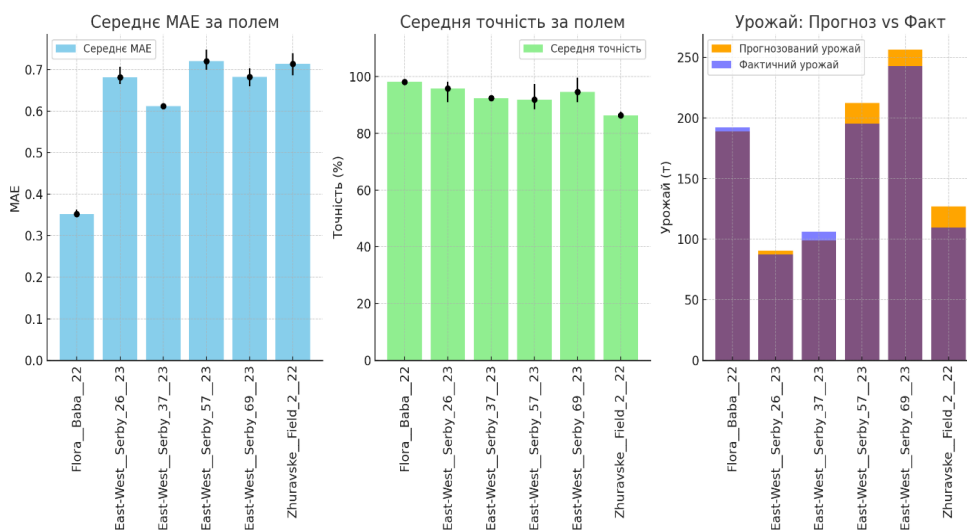


Рис. 3. Візуалізація отриманих показників точності.

Використання створеної моделі дозволило оцінити потенційну врожайність з досить високою точністю, як це видно з отриманих даних. Проведений аналіз показує, що модель не є достатньо стабільною: найнижчу точність зафіксовано на рівні 87.62%, а найвищу — на рівні 97.88%. Середнє значення помилки по всіх ділянках поля (MAE) становить 0.608. Середня точність прогнозування сумарної врожайності складає 92.78%.

Існує проблема вимушеного використання обмеженої вибірки для навчання та тестування. У навчальній вибірці наявні лише 8 полів, що значно обмежує здатність моделі узагальнювати знання та адаптуватися до нових умов. Це контрастує з сучасними методами прогнозування врожайності, де для навчання використовуються сотні, а інколи й тисячі полів, дозволяючи моделям краще розрізняти різноманітні патерни росту рослин. З розширенням навчальної вибірки, модель зможе ідентифікувати нові закономірності та покращити свою точність, стаючи більш ефективною у прогнозуванні врожайності в різноманітних агрокліматичних умовах та регіонах. Це важливо для підвищення точності

прогнозів і зробіть модель більш універсальною та придатною для широкого спектра застосувань.

На рисунках 4 та 5 наведено приклади візуалізації прогнозів. На кожному рисунку ліворуч розміщено карту реальної врожайності, праворуч — карту прогнозованої врожайності.

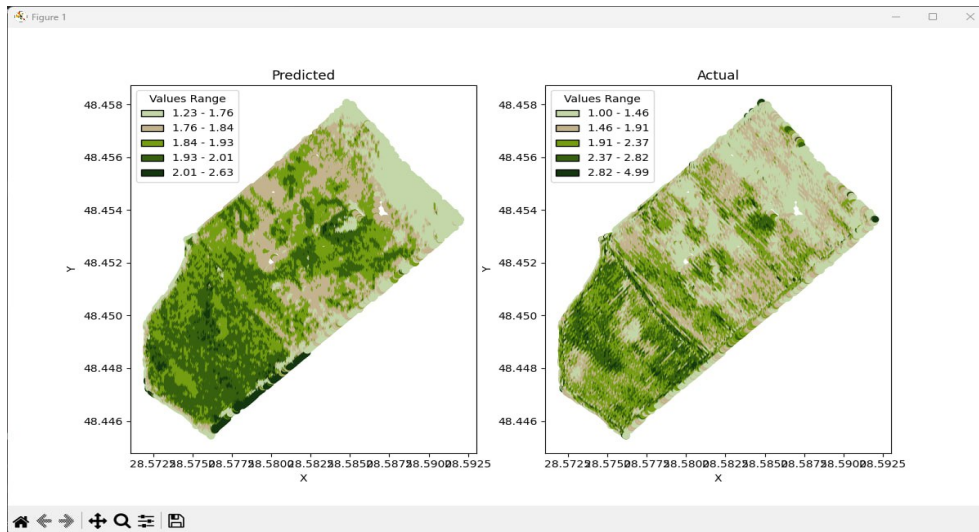


Рис. 4. Візуалізація результатів роботи моделі прогнозу, приклад 1.

**5. Висновки та перспективи подальших досліджень.** Завдяки впровадженню моделі, яка здатна виконувати деталізований прогноз з локалізацією до окремих ділянок, відкриваються нові перспективи застосування в сільськогосподарській індустрії. Такий підхід дозволяє проводити глибокий аналіз впливу локальних факторів на розвиток та врожайність рослин, що сприяє виявленню оптимальних умов для їхнього зростання. Крім того, це відкриває можливості для оптимізації диференційованого внесення добрив та хімічних речовин для захисту рослин, значно підвищуючи ефективність сільськогосподарських практик та мінімізуючи негативний вплив людської діяльності на довкілля.

Висока точність прогнозів, забезпечена застосуванням моделі, дозволяє з великою достовірністю здійснювати планування бюджету фермерських господарств. Це, в свою чергу, дозволяє агровиробникам ефективно планувати та оптимізувати витрати, забезпечуючи краще управління ресурсами та підвищення загальної рентабельності. Таким чином, реалізація описаної моделі відкриває значні можливості для підвищення продуктивності та підтримки сталого розвитку в аграрному секторі.

При подальших дослідженнях збагачення навчальної вибірки через додавання даних з різноманітних географічних локацій та умов вирощування стане ключем до максимізації універсальності моделі, дозволяючи їй ефективно працювати з ширшим спектром агроєкосистем.

Оптимізація навчальної вибірки шляхом балансування її структури видаленням надмірно представлених даних і аугментацією (збільшення обсягу) менш представлених категорій відіграє важливу роль у підвищенні точності прогнозів. Це дозволить моделі краще адаптуватися до варіативності умов та особливостей різних видів сільськогосподарських культур.

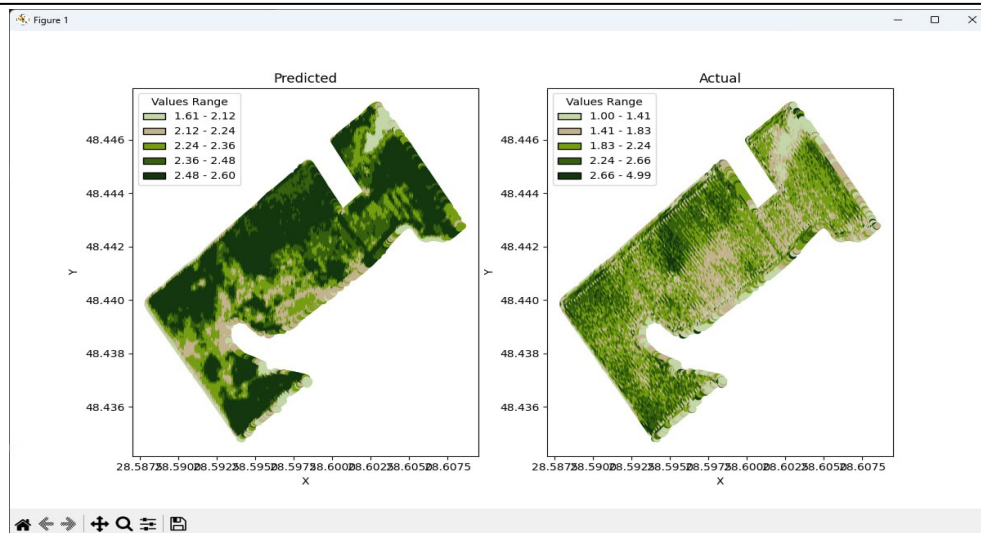


Рис. 5. Візуалізація результатів роботи моделі прогнозу, приклад 2.

Розробка методу для генерації та аналізу комбінацій можливих варіантів догляду за рослинами відкриває нові перспективи для створення комплексних систем рекомендацій. Це не тільки підвищить точність індивідуалізованих прогнозів для кожної ділянки, але й допоможе визначити найефективніші стратегії агротехнічних заходів, заснованих на глибокому аналізі потенційної врожайності.

Розширення сфери застосування моделі шляхом проведення досліджень на нових видах культур, виходячи за межі одного соняшнику, сприятиме генералізації та адаптації системи прогнозування для широкого спектру аграрних потреб. Такий підхід не тільки розширить можливості використання штучного інтелекту в агрономії, але й забезпечить більш глибоке розуміння впливу різноманітних факторів на продуктивність культур, сприяючи сталому розвитку аграрного сектору.

### Список використаної літератури

1. Khaki S., Wang L. Crop yield prediction using deep neural networks. *Front. Plant Sci.* 2019. Vol. 10. P. 621.
2. Paudel D., Boogaard H., A. de Wit, Janssen S., Osinga S., Pylianidis C., Athanasiadis I. N. Machine learning for large-scale crop yield forecasting. *Agric. Syst.* 2021. Vol. 187. P. 103016. DOI: <https://doi.org/10.1016/j.agsy.2020.103016>
3. Elavarasan D., Vincent P. M. D. Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. *IEEE Access.* 2020. Vol. 8. P. 86886–86901. DOI: <https://doi.org/10.1109/ACCESS.2020.2992480>
4. Al-Gaadi K. A., Hassaballa A. A., Tola E., Kayad A. G., Madugundu R., Alblewi B., Assiri F. Prediction of Potato Crop Yield Using Precision Agriculture Techniques. *PLoS One.* 2016. Vol. 11, No. 9. DOI: <https://doi.org/10.1371/journal.pone.0162219>
5. Зозуля О. Л. Сучасні методи цифрового моніторингу в рослинництві : Монографія / О. Л. Зозуля, В. В. Швартау, Л. М. Михальська, О. Л. Ковель, Г. М. Гнатієнко, В. Є. Снитюк, В. М. Домрачев, Н. П. Тменова. Київ : Від А до Я, 2023. 254 с.
6. Anusha P. V., Anuradha C., Murty P. S. R. C., Kiran C. S. Detecting Outliers in High Dimensional Data Sets using Z-Score Methodology. *International Journal of Innovative Technology and Exploring Engineering.* 2019. Vol. 9, No. 1. P. 48–53. DOI: <https://doi.org/10.35940/ijitee.A3910.119119>
7. Jiao L., Huo L., Hu C., and Tang P. Refined Unet: Unet-Based Refinement Network for

- Cloud and Shadow Precise Segmentation. *Remote Sensing*. 2020. Vol. 12, No. 12: 2001. DOI: <https://doi.org/10.3390/rs12122001>
8. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., and Liu T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. *31st Conference on Neural Information Processing Systems (NIPS 2017)* : Long Beach. CA : USA. 2017. Vol. 30. P. 3146–3154.
  9. Bilan S., Hnatiienko V., Ilarionov O., and Krasovska H. The Technology of Selection and Recognition of Information Objects on Images of the Earth's Surface Based on Multi-Projection Analysis. *CEUR Workshop Proceedings. Selected Papers of the III International Scientific Symposium "Intelligent Solutions" (IntSol-2023)* : Symposium Proceedings Kyiv — Uzhhorod : Ukraine. 2023 September 27–28. Vol. 3538. P. 23–32.
  10. Hnatiienko H., Domrachev V., Saiko V. Monitoring the condition of agricultural crops based on the use of clustering methods. *15th International Conference Monitoring of Geological Processes and Ecological Condition of the Environment* : Monitoring. Nov 2021. Vol. 2021. P. 1–5. DOI: <https://doi.org/10.3997/2214-4609.20215K2049>

**Hnatiienko V. H., Hnatiienko H. M., Zozulia O. L., Snytyuk V. Ye.** Method of forecasting yield of agricultural crops using multifactor analysis and neural networks.

The study focuses on the development of an intelligent yield forecasting system using satellite data, geoinformation data and climate indicators. The introduction of modern information technologies, such as machine learning and big data analytics, provides the agricultural sector with strategic advantages, reducing the risks of excessive use of desiccants and promoting sustainable agricultural development. This study is relevant for the agricultural sector, as it improves the accuracy of forecasts and the effectiveness of management decisions, while reducing costs and increasing the productivity of field processing.

**Keywords:** agricultural crops, desiccation, forecasting, satellite data, climate indicators, vegetation indices, neural networks.

## References

1. Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Front. Plant Sci.*, 10, 621.
2. Paudel, D., Boogaard, H., A. de Wit, Janssen, S., Osinga, S., Pylaniadis, C., & Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. *Agric. Syst.*, 187, 103016. <https://doi.org/10.1016/j.agsy.2020.103016>
3. Elavarasan, D., & Vincent, P. M. D. (2020). Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. *IEEE Access*, 8, 86886–86901. <https://doi.org/10.1109/ACCESS.2020.2992480>
4. Al-Gaadi, K. A., Hassaballa, A. A., Tola, E., Kayad, A. G., Madugundu, R., Alblewi, B., & Assiri, F. (2016). Prediction of Potato Crop Yield Using Precision Agriculture Techniques. *PLoS One*, 11(9). <https://doi.org/10.1371/journal.pone.0162219>
5. Zozulia, O. L. (2023). *Modern methods of digital monitoring in crop production: Monograph*. Kyiv: Vid A do Ya [in Ukrainian].
6. Anusha, P. V., Anuradha, C., Murty, P. S. R. C., & Kiran, C. S. (2019). Detecting Outliers in High Dimensional Data Sets using Z-Score Methodology. *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 48–53. <https://doi.org/10.35940/ijitee.A3910.119119>
7. Jiao, L., Huo, L., Hu, C., & Tang, P. (2020). Refined Unet: Unet-Based Refinement Network for Cloud and Shadow Precise Segmentation. *Remote Sensing*, 12(12), 2001. <https://doi.org/10.3390/rs12122001>
8. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach. CA: USA.
9. Bilan, S., Hnatiienko, V., Ilarionov, O., & Krasovska, H. (2023). The Technology of Selection and Recognition of Information Objects on Images of the Earth's Surface Based on Multi-

Projection Analysis. *CEUR Workshop Proceedings. Selected Papers of the III International Scientific Symposium "Intelligent Solutions" (IntSol-2023)*. Symposium Proceedings Kyiv — Uzhhorod: Ukraine.

10. Hnatiienko, H., Domrachev, V., & Saiko, V. (2021). Monitoring the condition of agricultural crops based on the use of clustering methods. *15th International Conference Monitoring of Geological Processes and Ecological Condition of the Environment*. Monitoring. <https://doi.org/10.3997/2214-4609.20215K2049>

Одержано 27.04.2024