

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВЕТЕРИНАРНОЇ
МЕДИЦИНИ ТА БІОТЕХНОЛОГІЙ ІМ. С.З.ГЖИЦЬКОГО
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

КВАЛІФІКАЦІЙНА РОБОТА

другого (магістерського) рівня вищої освіти

**на тему: «Інформаційна система прогнозування врожайності
озимої пшениці з використанням супутникових знімків та
глибинних нейронних мереж»**

Виконав: студент групи Іт-62

Спеціальності 126 «Інформаційні системи та
технології»

(шифр і назва)

Тирпак Віталій Мар'янович

(Прізвище та ініціали)

Керівник: к.т.н., в.о. доцента Ковалишин О.С.
(Прізвище та ініціали)

Рецензент: к.т.н., доцент Сиротюк С.В.
(Прізвище та ініціали)

ДУБЛЯНИ-2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ПРИРОДОКОРИСТУВАННЯ
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Другий (магістерський) рівень вищої освіти
Спеціальність 126 «Інформаційні системи та технології»

«ЗАТВЕРДЖУЮ»

Завідувач кафедри _____

д.т.н., проф. А.М. Тригуба

«_____» _____ 2025 р.

ЗАВДАННЯ

на кваліфікаційну роботу студенту

Тирпаку Віталію Мар'яновичу

1. Тема роботи: «Інформаційна система прогнозування врожайності озимої пшениці з використанням супутникових знімків та глибинних нейронних мереж»

Керівник роботи Ковашин Олег Степанович, доцент

затверджені наказом по університету від 28.02.2025 року № 140/к-с.

2. Строк подання студентом роботи 10.12.2025 р.

3. Вихідні дані до роботи: дані супутникового моніторингу стану посівів озимої пшениці; статистичні показники врожайності за попередні роки; метеорологічні параметри; карти землекористування та ґрунтових характеристик; архітектурні параметри глибинних нейронних мереж.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які необхідно розробити) _____

Вступ.

Аналіз стану прогнозування врожайності озимої пшениці та доцільності використання супутникових знімків та глибинних нейронних мереж.

Теоретичні основи та структура інформаційної системи прогнозування врожайності.

Розроблення моделі та реалізація програмного прототипу інформаційної системи прогнозування врожайності.

Охорона праці та безпека у надзвичайних ситуаціях.

Визначення економічної ефективності.

Висновки та пропозиції.

Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових слайдів): аналіз стану прогнозування врожайності озимої пшениці та доцільності використання супутникових знімків та глибинних нейронних мереж; теоретичні основи та структура інформаційної системи прогнозування врожайності; розроблення моделі та реалізація програмного прототипу інформаційної системи прогнозування врожайності; економічна ефективність.

6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1, 2, 3, 5	<i>Ковалишин О.С., в.о. доцента. кафедри інформаційних технологій</i>		
4	<i>Городецький І.М., доцент кафедри інженерної механіки</i>		

7. Дата видачі завдання

28 лютого 2025 р.

Календарний план

№ з/п	Назва етапів кваліфікаційної роботи	Терміни виконання етапів роботи	Примітка
1	<i>Написання першого розділу</i>	28.02-20.03.25	
2	<i>Виконання другого розділу та аркушів ілюстраційного матеріалу до нього</i>	21.03-14.06.25	
3.	<i>Виконання третього розділу та аркушів ілюстраційного матеріалу до нього</i>	15.06.25-10.07.25	
4.	<i>Написання розділу «Охорона праці та безпека у надзвичайних ситуаціях»</i>	11.07-31.08.25	
5.	<i>Оцінення ефективності запропонованої системи</i>	01.09-31.10.25	
6.	<i>Завершення оформлення розрахунково-пояснювальної записки та аркушів ілюстраційного матеріалу</i>	01-30.11.25	
7.	<i>Завершення роботи в цілому</i>	01-10.12.25	

Студент _____ Тирпак В.М.
(підпис)

Керівник роботи _____ Ковалишин О.С.
(підпис)

УДК 004.891.3:004.932.72:631.559

Інформаційна система прогнозування врожайності озимої пшениці з використанням супутникових знімків та глибинних нейронних мереж.

Тирпак В.М. Кафедра інформаційних технологій – Дубляни, ЛНУВМ, 2025.

Кваліфікаційна робота: 89 с. текст. част., 18 рис., 8 табл., 15 арк. ілюстраційного матеріалу, 33 джерела.

Подано особливості та доцільність використання супутникових знімків і глибинних нейронних мереж для прогнозування врожайності озимої пшениці. Проаналізовано сучасні тенденції в агромоніторингу та методах прогнозування. Розглянуто можливості застосування даних дистанційного зондування Землі для підвищення точності оцінки стану посівів і прогнозування врожайності. Сформульовано мету, завдання та етапи виконання кваліфікаційної роботи.

Розроблено теоретичні основи та структуру інформаційної системи прогнозування врожайності озимої пшениці з використанням архітектур глибинного навчання. Проведено порівняльний аналіз моделей CNN, LSTM, GRU, гібридних CNN–LSTM та Transformer, визначено їх ефективність для задач обробки часових рядів аграрних даних. Обґрунтовано вибір інструментальних засобів і програмного середовища для реалізації системи, що поєднує супутникові знімки, метеопоказники та алгоритми глибинного навчання.

Створено програмний прототип інформаційної системи прогнозування врожайності, що виконує збір, попередню обробку та аналіз даних дистанційного зондування, побудову й навчання нейромережевої моделі. Наведено результати оцінки точності прогнозів і ефективності моделі. Побудовано структурну схему та логічну організацію системи, які забезпечують інтеграцію модулів аналізу, прогнозування та візуалізації результатів.

Виконано аналіз умов праці під час монтажу інформаційної системи. Проведено розрахунок економічної ефективності.

Ключові слова: прогнозування врожайності, озима пшениця, глибинне навчання, супутникові знімки, дистанційне зондування Землі, інформаційна система.

ЗМІСТ

ВСТУП.....	7
Розділ 1.....	9
Розділ 1. АНАЛІЗ СТАНУ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ОЗИМОЇ ПШЕНИЦІ ТА ДОЦІЛЬНОСТІ ВИКОРИСТАННЯМ СУПУТНИКОВИХ ЗНІМКІВ ТА ГЛИБИННИХ НЕЙРОННИХ МЕРЕЖ.....	9
1.1. Сучасні тенденції у прогнозуванні врожайності зернових культур.....	9
1.2. Особливості озимої пшениці як об'єкта прогнозування	15
1.3. Методи та моделі прогнозування врожайності	18
1.4. Використання супутникових даних і дистанційного зондування Землі у сільському господарстві.....	22
1.5. Завдання кваліфікаційної роботи.....	25
Розділ 2. ТЕОРЕТИЧНІ ОСНОВИ ТА СТРУКТУРА ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ	27
2.1. Теоретичні засади прогнозування врожайності сільськогосподарських культур.....	27
2.2. Особливості використання супутникових знімків у задачах агромоніторингу	31
2.3. Архітектури глибинних нейронних мереж для прогнозування часових рядів.....	35
2.3.1. Архітектура згорткової нейронної мережі (CNN)	36
2.3.2. Архітектура рекурентної нейронної мережі LSTM.....	37
2.3.3. Архітектура GRU (Gated Recurrent Unit)	38
2.3.4. Гібридна архітектура CNN–LSTM.....	39
2.3.5. Архітектура Transformer.....	40
2.4. Вибір інструментальних та програмних засобів для реалізації системи	42
Розділ 3. РОЗРОБЛЕННЯ МОДЕЛІ ТА РЕАЛІЗАЦІЯ ПРОГРАМНОГО ПРОТОТИПУ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ	46
3.1. Збір, підготовка та аналіз вхідних даних дистанційного зондування і метеопараметрів.....	46
3.2. Побудова та навчання глибинної нейронної моделі прогнозування врожайності	52

3.3. Оцінка ефективності та точності побудованої моделі	55
3.4. Структурна схема та логічна організація інформаційної системи прогнозування	58
Розділ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	62
4.1. Аналіз умов праці та заходи із їх покращення.....	62
4.2. Розробка логічно-імітаційної моделі травматизму під час монтажу інформаційної системи прогнозування врожайності	65
4.3. Розробка заходів щодо безпеки у надзвичайних ситуаціях	68
Розділ 5. ВИЗНАЧЕННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ОЗИМОЇ ПШЕНИЦІ	70
ВИСНОВКИ І ПРОПОЗИЦІЇ.....	73
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	76
ДОДАТКИ.....	80
Додаток А.1. Код підготовки вхідних даних дистанційного зондування і метеопараметрів.....	81
Додаток А.2. Код системи прогнозування врожайності	84

ВСТУП

У сучасних умовах глобального потепління, частих проявів кліматичної нестабільності та зростаючого тиску на продовольчу безпеку прогнозування врожайності стає критично важливим інструментом для управління аграрним виробництвом. Озима пшениця – одна з найважливіших зернових культур України, що потребує точних методів оцінки майбутніх врожаїв для планування аграрної політики, оптимального розподілу ресурсів і мінімізації ризиків втрат [6].

Традиційні моделі прогнозування базуються на статистичних підходах (методи множинної регресії, кореляційний аналіз) з застосуванням метеорологічних, агрохімічних та історичних даних врожайності. Проте ці методи часто обмежені лінійністю зв'язків, просторовою агрегацією і слабкою адаптацією до нових умов. У той же час дистанційне зондування Землі, зокрема супутникові знімки середньої та високої роздільної здатності, дає змогу оперативно отримувати інформацію про стан рослинного покриву протягом вегетаційного періоду та відслідковувати просторово-часові зміни індексів рослинності (наприклад NDVI, EVI) [7].

Застосування глибинного навчання (deep learning) дозволяє виявляти складні нелінійні взаємозв'язки між спектральними показниками супутникових зображень, метеорологічними факторами і фактичною врожайністю. У ряді досліджень показано, що моделі на основі CNN, LSTM або їх гібридних варіантів демонструють кращу прогностичну здатність порівняно з класичними машинними методами – хоча залишаються проблеми інтерпретованості та необхідності великої кількості даних [8].

В Україні були впроваджені спроби прогнозування врожайності озимої пшениці з використанням супутникових даних: наприклад, в дослідженні [9] та співавторів показано можливість використання NDVI даних із супутника в поєднанні з метеорологічними показниками для прогностичних моделей на рівні областей.

Таким чином, поєднання супутникових даних із глибинними нейронними мережами має значний потенціал для підвищення точності прогнозів урожайності. Водночас реалізація таких систем вимагає врахування специфіки місцевих умов, обробки даних, вибору архітектур, оптимізації параметрів моделей і їхньої валідації в українських агрокліматичних умовах.

Об'єктом дослідження є процес прогнозування врожайності озимої пшениці у сільськогосподарському виробництві.

Предметом дослідження стають методи обробки супутникових зображень та архітектури глибинних нейронних мереж як інструменти побудови моделей прогнозування врожайності.

Засобами дослідження у роботі використовуватимуться супутникові знімки (наприклад Sentinel-2, Landsat, MODIS), набори метеорологічних даних, агрономічні дані врожайності, а також алгоритми глибинного навчання (CNN, LSTM, гібриди), інструменти обробки та попередньої обробки даних (корекція хмарності, нормалізація спектральних каналів, створення вегетаційних індексів тощо).

Метою цієї роботи є розроблення інформаційної системи, здатної автоматично прогнозувати врожайність озимої пшениці з використанням супутникових даних дистанційного зондування Землі та моделей глибинного навчання, що поєднують у собі оптимізаційні алгоритми й сучасні підходи до обробки часових рядів вегетаційних індексів.

Такий підхід спрямований на підвищення точності прогнозів у контексті адаптації аграрного виробництва до кліматичних змін, оптимізації використання ресурсів і формування ефективної політики продовольчої безпеки.

Розділ 1.

АНАЛІЗ СТАНУ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ОЗИМОЇ ПШЕНИЦІ ТА ДОЦІЛЬНОСТІ ВИКОРИСТАННЯМ СУПУТНИКОВИХ ЗНІМКІВ ТА ГЛИБИННИХ НЕЙРОННИХ МЕРЕЖ

1.1. Сучасні тенденції у прогнозуванні врожайності зернових культур

Упродовж останніх років у наукових дослідженнях з прогнозування врожайності зернових культур спостерігається значне зрушення від класичних статистичних і регресійних методів до методів із застосуванням машинного та глибинного навчання. Це зумовлено швидким збільшенням обсягів доступних даних (метеорологічних, агрономічних, супутникових), підвищенням обчислювальних можливостей, та розвитком технологій дистанційного зондування. Головна прагнення – підвищити точність прогнозів на ранніх етапах вегетації, мінімізуючи затрати часу і ресурсів.

Класичні підходи, такі як множинна регресія, методи часових рядів або кореляційні моделі (наприклад, регресія з факторами температури, опадів, агрохімічними показниками), залишаються основою у багатьох роботах. Проте їхня здатність моделювати складні нелінійні взаємозв'язки обмежена. З розвитком дистанційного зондування, особливо з використанням оптичних і радарних супутників (Sentinel, Landsat, MODIS), дослідники отримали змогу інтегрувати часові серії спектральних індексів (NDVI, EVI, SAVI та ін.), показників вологості ґрунту, температур та інші змінні, що значно підсилює прогностичний потенціал моделей [10].

Сучасні тенденції в розвитку прогнозних систем можна класифікувати за кількома напрямками, які описано нижче на рис. 1.1.



Рисунок 1.1 – Напрямки розвитку прогнозних систем

Застосування глибоких нейронних мереж (Deep Learning)

Серед найбільш популярних архітектур – згорткові нейронні мережі (CNN), рекурентні мережі (LSTM, GRU) або їхні гібридні поєднання (наприклад, CNN–RNN). Такі моделі здатні враховувати просторово-часову структуру даних: CNN витягує ознаки з растрових шарів спектральних зображень, а RNN/ LSTM обробляє часові залежності між вегетаційними фазами [11].

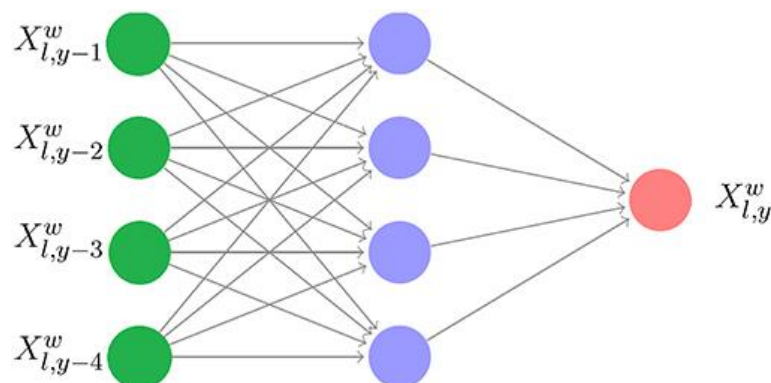


Рисунок 1.2 – Структура нейронної мережі для прогнозування погоди [11]

Наприклад, гібридна модель CNN–RNN показала значно кращу точність прогнозу врожайності кукурудзи і сої на базі кліматичних та ґрунтових даних у

порівнянні з класичними підходами (RMSE становив ~9% від середнього врожаю) [12].

Мультимодальні дані та їх злиття (data fusion)

Для підвищення продуктивності моделей дослідники поєднують різноманітні джерела інформації: супутникові зображення, метеодані, ґрунтові характеристики, агрономічні дані управління (наприклад, удобрення, сівоzmіна) [5]. Злиття таких даних дозволяє охопити як просторові, так і екологічні фактори, що впливають на продуктивність. У випадку моделей зі злиттям даних часто використовують механізми уваги (attention), або багатоголосні (multi-head) стратегії, які керують важливістю кожного джерела даних [14].

Підхід трансферного навчання (Transfer Learning) та «слабке навчання»

(Weakly Supervised Learning)

Оскільки для багатьох регіонів бракує достатньої кількості міток (даних урожаїв) для тренування потужних моделей, застосовують трансферні моделі – навчають мережі на регіонах з багатьма даними, а потім адаптують на регіони з обмеженими даними. Також з'являються підходи weak supervision, за яких модель навчається на низькороздільних мітках, але видає прогнози високого просторового оцінювання [7]. Такий підхід успішно застосовано, наприклад, для озимої пшениці у Європі, коли макрорегіональні дані про врожай були використані для побудови високороздільного прогнозу на підрегіони [15].

Підвищення інтерпретованості моделей (Explainable AI)

Одна з ключових проблем глибокого навчання – «чорний ящик». У відповідь на це дослідники розробляють інтерпретовані варіанти моделей, або обчислюють значущість ознак у часовому контексті. Наприклад, методи на кшталт SHAP, LIME або вбудовані модулі уваги дозволяють встановити, які періоди або спектральні канали найбільше впливають на прогноз [9]. Такий підхід особливо важливий в аграрному секторі, де менеджери і агрономи повинні довіряти результатам моделі.

Розробка моделей в реальному часі та раннє прогнозування

Ще одна тенденція – створення моделей, які можуть робити прогноз урожайності на ранніх стадіях вегетації (наприклад, після першого або другого місяця посіву). Це дає змогу фермерам коригувати агробіологічні заходи. Однак для цього доводиться працювати з урізаними часовими серіями, що підвищує ризик помилки. На практиці дослідження показують, що повна вегетаційна серія (кілька часових кадрів) дає кращий прогноз, ніж ранні прогнози, але гібридні моделі з attention дозволяють знизити похибку навіть на коротких серіях [16].

Таблиця 1.1 – Узагальнена структура моделі прогнозування

Етап	Опис
Збір даних	Супутникові багатоспектральні кадри, метеодані, агрономічні показники
Попередня обробка	Вирізання хмарності, нормалізація, інтерполяція відсутніх даних
Генерація ознак	Обчислення індексів NDVI, EVI, індексів вологості, статистичних текстур
Побудова моделі	Архітектура (CNN, LSTM, гібриди, attention)
Навчання та валідація	Розподіл даних на тренувальний і тестовий набори, перехресна валідація
Інтерпретація	Визначення впливу ознак за допомогою SHAP, LIME, attention-мап
Прогнозування	Отримання карти врожайності в пік вегетації або перед збиранням

Застосування дронів (UAV) і високороздільного зондування

Крім супутникових даних, дедалі популярнішими стають знімки з безпілотників, що забезпечують високу просторову роздільність на рівні ділянок чи навіть рослини. У роботах показано, що комбінація UAV-зображень із супутниковими даними збільшує точність прогнозу – особливо у польових

дослідах з детальною деталізацією [17]. Такі дані особливо зручні для виявлення просторової неоднорідності посівних площ.

Оцінка якості та пошук меж масштабованості

У літературі триває дискусія щодо обмежень моделей: здатність узагальнення для нових регіонів, стабільність у змінних кліматичних умовах, чутливість до шуму даних. Деякі дослідження показали, що моделі з хорошими результатами в одному регіоні не завжди коректно працюють в іншому через відмінність ґрунтово-кліматичних умов [11]. Також звертають увагу на стратегії зменшення надмірного підлаштування (regularization), обмеженням складності мереж і валідацію на незалежних регіонах.

Нижче наведена спрощена математична модель, яка ілюструє, як у тренуванні глибинної мережі враховуються часові та просторові компоненти:

$$\hat{Y}_{i,t} = f(X_{i,t-k:t}, M_i, P_i), \quad (1.1)$$

де $\hat{Y}_{i,t}$ – прогнозована врожайність (для ділянки i на час t); $X_{i,t-k:t}$ – часовий ряд мультиспектральних і метеофакторних даних за період k кадрів до моменту t ; M_i – постійні характеристики ділянки (ґрунтові властивості, топографія); P_i – фактори управління (дозування добрив, сівозміна тощо).

У цій формулі функція f може реалізовуватись у вигляді глибокої нейронної мережі типу CNN-LSTM або трансформера. Оптимізація відбувається шляхом мінімізації середньоквадратичної або середньоабсолютної похибки:

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N (Y_{i,t} - \hat{Y}_{i,t})^2 + \lambda \cdot \Omega(\theta), \quad (1.2)$$

де $Y_{i,t}$ – фактична врожайність; θ – параметри моделі; $\Omega(\theta)$ – регуляризатор; λ – коефіцієнт регуляризації.

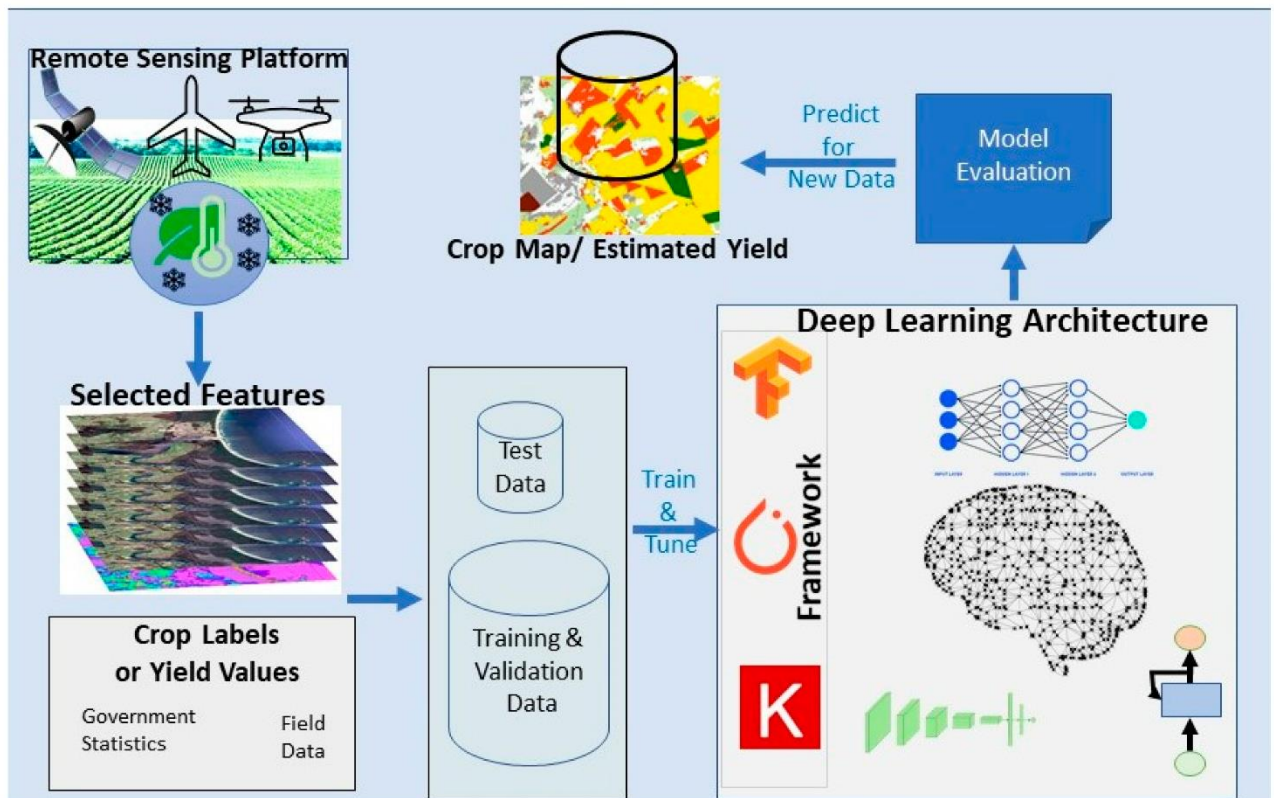


Рисунок 1.3 – Типовий підхід у дослідженнях RS + DL [18]

На рисунку 1.4 зверху (ілюстрація із джерела) показано типовий підхід у дослідженнях RS + DL: як дані від супутника або інших джерел передаються до шару ознакових карт і потім входять у глибинні моделі з виходом у вигляді карти врожайності або оцінок агрономічних параметрів.

Ці тенденції відображають поступову еволюцію від моделювання лінійних залежностей до інтелектуальних систем, здатних навчатися із великого обсягу даних і враховувати взаємодії багатьох факторів. Проте ці підходи все ще мають виклики – зокрема, потреба у верифікації моделі в нових регіонах, забезпечення інтерпретованості для кінцевих користувачів, а також питання з якості даних і обробки шуму. Надалі – важливо не просто збільшувати складність моделей, а працювати над їхньою надійністю, масштабованістю та прозорістю.

1.2. Особливості озимої пшениці як об'єкта прогнозування

Озима пшениця вирізняється серед зернових культур специфічним циклом розвитку, що включає фазу перезимівлі, весняне відновлення росту і літнє досягання. Саме ці особливості створюють певні виклики й можливості для побудови прогностичних моделей. По-перше, проходження зимівлі, холодостійкість (зимостійкість) і здатність до відновлення вегетації весною – це важливі біологічні властивості, які суттєво впливають на кінцеву врожайність. У сучасних сортах озимої пшениці велика увага приділяється врожайності за умов змінного клімату, де здатність пережити зиму може стати обмежуючим фактором [1].

У працях, присвячених адаптаційним характеристикам сучасних сортів озимої пшениці в Українському Лісостепу, показано, що тривалість верналізаційного періоду (необхідного для переходу рослини з вегетативного стану до генеративного) може відрізнятися між сортами – деякі сорти мають коротку верналізацію, інші – середню. Фотоперіодична чутливість також варіює: сорти з помірною чутливістю до тривалості дня менш схильні до коливань у погодних умовах навесні, що дає перевагу у прогнозуванні їхньої продуктивності [1].

Крім того, коренева система озимої пшениці часто має глибоке проникнення в ґрунт (до 1,5–2 м за сприятливих умов), що дозволяє використати підґрунтову вологу в посушливі періоди. Це створює більш складну динаміку водозабезпечення рослин: поверхнева вологість ґрунту може виявитися менш критичною, ніж глибока волога, але обидві обумовлюють сумарні водні ресурси, які можуть враховуватися в моделях. Коренева інтенсивність, розгалуженість і ефективність водозабору – це важливі характеристики, які можуть бути опосередковано відображені через індекси вологи або мультиспектральні канали.

Вегетаційні фази озимої пшениці також мають чітку структуру: період осіннього росту, фаза куціння (утворення пагонів), зимівля, весняне

відновлення, стеблуння, цвітіння, налив зерна й дозрівання. Ці етапи можна узагальнити через фенологічні стадії, наприклад за шкалою BBCH або Zadoks, що дає змогу синхронізувати дані супутникових знімків з біологічними процесами [2]. Зміни індексів рослинності (NDVI, EVI тощо) та спектральні аномалії в різні фази дозволяють моделям “відчувати”, на якій стадії знаходиться культура, і коригувати прогноз залежно від часу.

Крім біологічного аспекту, агротехнічні чинники також відіграють виняткову роль: оптимальні строки сівби, попередник урожаю, глибина і норма сівби, живлення (особливо азотом) – усе це впливає на структуру рослин, листову поверхню, коренева маса та потенціал фотосинтезу. У дослідженнях показано, що пізні строки сівби можуть призвести до недостатньої підготовки до зимівлі, зменшення числа пагонів або меншої вегетативної маси, що погано корелює з кінцевим урожаєм [3].

Важливо зазначити, що вплив погодних умов (температура, опади, випаровування) на озиму пшеницю неоднаковий у різні фази. Наприклад, у фазу наливу зерна нестача вологи має більш драматичний ефект, ніж на ранніх фазах. У дослідженні Ma et al. (2025) встановлено, що серед змінних, які найсильніше впливають на врожайність озимої пшениці, – опади та вологість ґрунту, потім – температура та випаровування [4]. Це означає, що у моделях прогнозування важливо вважати не тільки середні величини, але й їхню часову динаміку та взаємодії між фазами.

З математичної точки зору, для озимої пшениці часто використовують функції, які враховують фазу часу та накопичені значення. Наприклад, можна уявити прогнозну модель у формі:

$$\hat{Y} = f\left(\sum_{t=1}^T w_t \cdot X_t, C\right), \quad (1.3)$$

де X_t – вектор ознак у t -му періоді (спектральні індекси, метеоумови, ґрунтові показники); w_t – ваговий коефіцієнт кожної фази (може бути модельованим як функція уваги); C – константні характеристики поля (ґрунт, топографія, сорт).

Таке представлення дозволяє врахувати, що однакові показники в різні фази матимуть неоднаковий вплив на врожай.

Щодо структури даних, для озимої пшениці моделі часто використовують часові ряди з інтервалом від тижня до декади. Наприклад, щотижневе підсумування індексів дозволяє згладити шум і краще вловити тренди. Також у багатьох роботах показано, що використання вікон (наприклад, ковзне середнє з трьома тижнями) дає кращі результати у стабільності прогнозу.

Складові моделі прогнозування озимої пшениці представлені у таблиці 1.2.

Таблиця 1.2 – Складові моделі прогнозування озимої пшениці

Компонент	Призначення
Вхідні дані	часові індекси, метеоумови, ґрунтові показники, сортові характеристики
Попередня обробка	фільтрація шуму, нормалізація, обрізка хмарності
Генерація ознак	ковзне середнє, лаги, інтеракції між ознаками
Модель прогнозу	глибинна мережа (CNN, RNN, трансформер тощо)
Інтерпретація	оцінка ваг фаз або ознак (attention, SHAP тощо)

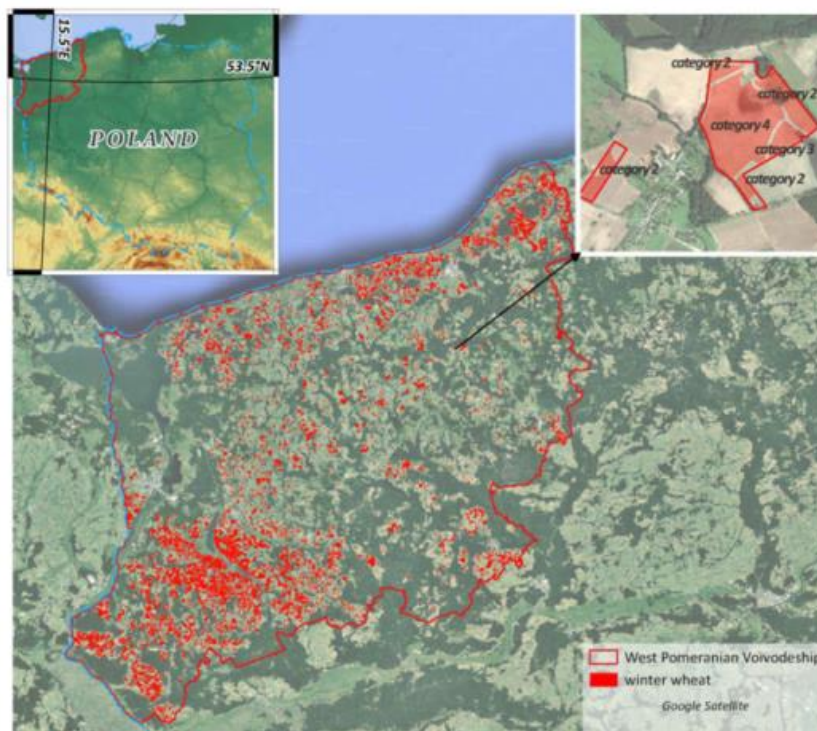


Рисунок 1.4 – Зміни індексу NDVI протягом вегетації [19]

Рисунок 1.4 демонструє, як у польових умовах змінюється індекс NDVI протягом вегетації (зростання, плато, спад) і як він корелює з основними етапами розвитку пшениці.

Цей набір особливостей – біологічні, агротехнічні, кліматичні – формує специфічний контекст для розробки моделей прогнозування врожайності озимої пшениці. Для високої точності моделей важливо не просто включати багато ознак, а коректно враховувати фазовість розвитку, кореневу динаміку й адаптивні характеристики сорту.

1.3. Методи та моделі прогнозування врожайності

Упродовж десятиліть прогнозування врожайності ґрунтувалося на статистичних підходах, які давали змогу проводити оцінку залежностей між метеорологічними факторами, агрономічними даними й історичними показниками врожаю. Класичні методи, такі як лінійна регресія, множинна регресія, кореляційний аналіз або методи часових рядів, залишаються фундаментом багатьох систем прогнозування. Наприклад, модель множинної лінійної регресії приймає форму:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon, \quad (1.4)$$

де Y – врожайність; $X_1, \dots, X_p$ – пояснювальні змінні (температура, опади, ґрунтові показники тощо), β_i – коефіцієнти регресії, ε – випадкова похибка.

Ці моделі зручні своєю прозорістю й простотою інтерпретації, але їхня слабка сторона – обмеження в адекватному моделюванні складних нелінійних взаємодій між факторами, особливо коли йдеться про часові залежності або взаємодії між агрокліматичними змінними [20].

Ще один класичний підхід – моделі часових рядів (ARIMA, SARIMA тощо), які мають змогу моделювати автокореляційні структури даних. Наприклад, моделі ARIMA можуть використовуватися для прогнозу

врожайності на основі тенденцій в минулих роках за умови, що сезонність та тренди є стабільними. Однак вони погано справляються з зовнішніми впливами – змінним кліматом, екстремальними погодними подіями чи інтенсивними агротехнічними змінами.

З розвитком агрономічного моделювання з'явилися процесно-орієнтовані (mechanistic або біофізичні) моделі, які описують фізіологічні процеси росту рослин – фотосинтез, транспірація, динаміка поглинання поживних речовин, вплив водного режиму. Такі моделі (наприклад, DSSAT, APSIM тощо) дозволяють змодельовати урожайність як наслідок комплексного впливу навколишнього середовища та агротехнічних заходів. Вони корисні тим, що дають інтерпретовані змінні (сумарна продукція біомаси, коефіцієнт використання води тощо), але вимагають великих обсягів вхідних даних та точного налаштування параметрів. Часто процесно-орієнтовані моделі комбінують із статистичними або машинними підходами в гібридні системи, щоб поєднати експліцитні знання з здатністю до навчання з даних.

У останнє десятиліття все більшу популярність здобули моделі машинного навчання (ML) та, зокрема, глибинного навчання (DL). Основна ідея – дозволити модель фактично «вчитися» на великій кількості даних, виявляючи складні закономірності, які важко формалізувати вручну. Усього лиш кілька ознак – температура, опади, індекси NDVI та інші – можуть комбінуватися з надзвичайно широким спектром взаємодій. У роботі-огляді показано, що методи ML, такі як Random Forest, Support Vector Regression, градієнтний бустинг і інші, вже значно поліпшили прогнози врожайності порівняно з класичною регресією [20].

Особливо помітне зростання – у застосуванні глибинних нейронних мереж. Архітектури, як згорткові (CNN), рекурентні (LSTM) або їхні гібриди, здатні враховувати як просторові, так і часові залежності одночасно. Наприклад, у роботі [11] запропоновано CNN-RNN мережу, яка обробляє часові ряди кліматичних та агрономічних даних, досягаючи відносної помилки менш ніж 9% середнього врожаю. У систематичному огляді зазначається, що

глибинні моделі демонструють кращу узагальнювальну здатність на нових ділянках та кращу адаптивність до несподіваних неоднорідностей.

Моделі машинного і глибинного навчання часто використовують просторово-часові дані (супутникові знімки) разом із зовнішніми змінними. У таких підходах застосовують часові вікна, лагові ознаки, інтеракційні терміни та механізми уваги (attention), щоб модель могла сама обирати, які часові періоди або спектральні канали мають більший вплив. Наприклад, в одному з досліджень запропоновано LSTM з механізмом уваги, який адаптивно зважує вагу кожного часу у виході моделі [14].

З погляду математичної побудови, гібридна модель може бути записана як:

$$\hat{Y} = g(h_{temporal}(X_{t-k:t}), c(X_{static})), \quad (1.5)$$

де $h_{temporal}(\cdot)$ – функція часової обробки (наприклад, LSTM); $c(\cdot)$ – функція обробки статичних даних (тип ґрунту, топографія); g – вихідний шар, який об'єднує ці частини.

Крім того, зараз з'являються моделі, що використовують трансформери (transformer-based architectures) для часових рядів у сільському господарстві. Вони дозволяють більш ефективно моделювати довгострокові залежності завдяки механізмам самоуваги (self-attention). Деякі роботи також комбінують трансформери з CNN, отримуючи потужні просторово-часові мережі.

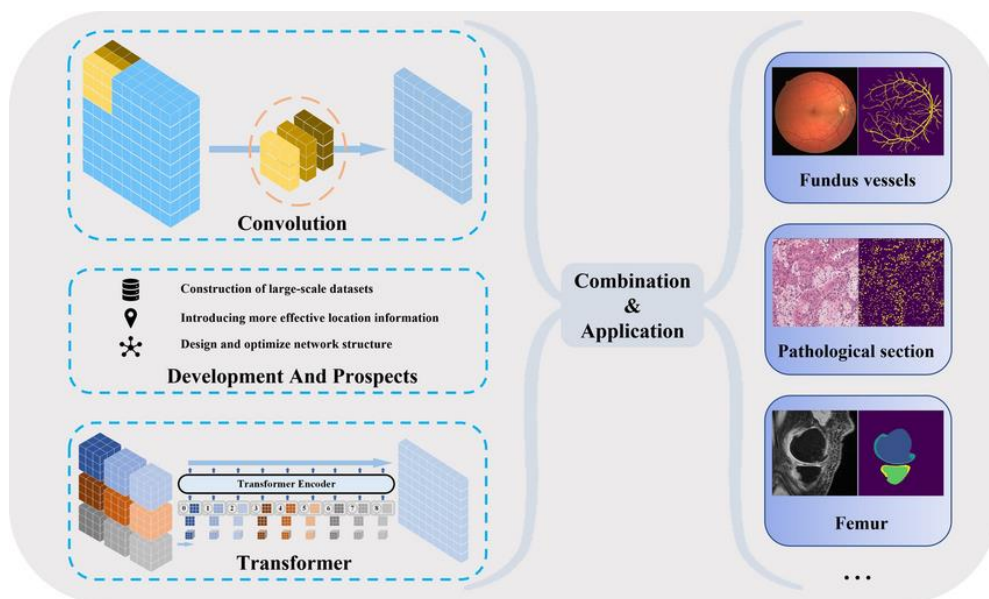


Рисунок 1.5 – Поєднання CNN з Transformer для завдання сегментації [21]

Не менш важливим напрямком є ансамбльні моделі – поєднання кількох підходів (регресія, дерево рішень, глибинна мережа) через вагове або більш складне усереднення. Ця стратегія дозволяє компенсувати слабкі місця окремих моделей, підвищуючи стабільність прогнозу у різних кліматичних умовах або агрокліматичних зонах.

Останнім часом з'являються комбіновані підходи, які поєднують глибоке навчання з процесно-орієнтованим моделюванням. Наприклад, деякі роботи пропонують використовувати результати біофізичних моделей як вхідні ознаки для нейронних мереж, щоб забезпечити додатковий шар екологічної інтерпретації. Це дозволяє скористатися «знаннями з природної моделі» та «гнучкістю навчальної моделі» одночасно.

Таблиця 1.3 – Характеристика методів та моделей прогнозування врожайності

Підхід	Сильні сторони	Обмеження
Статистичні (регресія, ARIMA)	Прозорість, простота інтерпретації	Нелінійність, обмежена адаптація
Процесні моделі	Біофізичні інтерпретовані змінні	Висока чутливість до точності вхідних даних
Машинне навчання (RF, SVR тощо)	Гнучкість, здатність враховувати нелінійності	Потреба у великому наборі даних, ризик перенавчання
Глибинне навчання (CNN, LSTM, трансформери)	Потужність, виявлення складних зв'язків	Чорний ящик, висока потреба в даних та обчислювальних ресурсах
Енсамблі	Стабільність, узагальнення	Складність налаштування, інтерпретація

Отже, сучасний ландшафт методів прогнозування врожайності – це поступовий перехід від простих статистичних моделей до високоефективних гібридних систем із елементами глибинного навчання, трансформерами,

механізмами уваги та ансамблями. Кожен підхід має свої переваги й обмеження, і вибір залежить від обсягу та якості даних, ресурсів, цілей прогнозу й масштабів застосування.

1.4. Використання супутникових даних і дистанційного зондування Землі у сільському господарстві

Використання супутникових технологій у сільському господарстві почалося ще у 1970-х роках із запуском програми Landsat, яка забезпечила багаторічні серії спостережень за рослинним покривом на глобальному рівні. Поступово дистанційне зондування стало невід'ємним інструментом агромоніторингу, адже воно дозволяє оцінювати стан посівів, прогнозувати врожайність, виявляти стресові ситуації та планувати використання ресурсів [22]. На відміну від класичних методів наземних обстежень, супутникові системи забезпечують регулярність і масштабність збору інформації, що особливо важливо в умовах зміни клімату.

Сьогодні у сільському господарстві застосовуються різні типи супутникових сенсорів. Гіперспектральні камери дозволяють аналізувати десятки вузьких спектральних діапазонів, що підвищує здатність до виявлення хвороб, дефіциту поживних речовин та диференціації культур. Термальні сенсори дають змогу оцінювати температуру поверхні ґрунту і рослин, що корисно для визначення випаровування та водного балансу. Радарні системи (SAR) особливо цінні у регіонах із високою хмарністю, оскільки вони працюють незалежно від погодних умов, надаючи інформацію про структуру рослинного покриву та вологість ґрунту [23].

Важливим завданням сучасного аграрного моніторингу є оцінка впливу кліматичних ризиків, таких як посуха. Для цього використовуються спеціалізовані індекси, зокрема Vegetation Health Index (VHI), який об'єднує температурні дані та вегетаційні індекси:

$$VHI = \alpha \cdot VCI + (1 - \alpha) \cdot TCI, \quad (1.6)$$

де VCI – індекс стану рослинності; TCI – температурний індекс; α – ваговий коефіцієнт, який визначає відносну значущість факторів.

Такий підхід дозволяє оцінити як фізіологічний стан культур, так і температурні стреси.

Крім окремих індексів, зростає значення інтеграції супутникових даних з геоінформаційними системами (ГІС). Це відкриває можливість створювати карти ризиків, просторово моделювати врожайність і виявляти зони, які потребують додаткового агротехнічного втручання. У багатьох країнах створено національні програми моніторингу сільського господарства на основі ДЗЗ, наприклад Crop Monitor у США чи системи раннього попередження про посухи в Африці. В Україні також реалізуються проєкти з використанням супутникових даних для прогнозування врожайності пшениці та кукурудзи [25].

Таблиця 1.4 – Основні напрями застосування різних сенсорів у сільському господарстві

Тип сенсора	Основні параметри	Застосування
Оптичні (Sentinel-2, Landsat)	Відбита енергія у видимому та ІЧ діапазонах	Вегетаційні індекси, класифікація культур
Радарні (Sentinel-1, RADARSAT)	Активне мікрохвильове випромінювання	Вологість ґрунту, структура рослинного покриву
Гіперспектральні (EnMAP, PRISMA)	До сотень вузьких спектральних каналів	Виявлення стресів, диференціація сортів
Термальні (MODIS, ECOSTRESS)	Температура поверхні	Випаровування, тепловий стрес культур

Технології машинного навчання все частіше поєднуються із супутниковими спостереженнями для створення аналітичних систем. На практиці інтеграція даних із Sentinel-1 та Sentinel-2 дає змогу одночасно

враховувати структурні та спектральні характеристики посівів. У дослідженні [26] показано, що поєднання оптичних та радарних даних значно покращує ідентифікацію фаз розвитку озимої пшениці під час посухи. Це відкриває нові перспективи для прогнозування врожайності з урахуванням ризиків, пов'язаних зі зміною клімату.

У таблиці 1.4 узагальнено основні напрями застосування різних сенсорів у сільському господарстві.

На рисунку 1.6 нижче показано приклад використання супутникових даних Sentinel-1 та Sentinel-2 для моніторингу озимої пшениці.

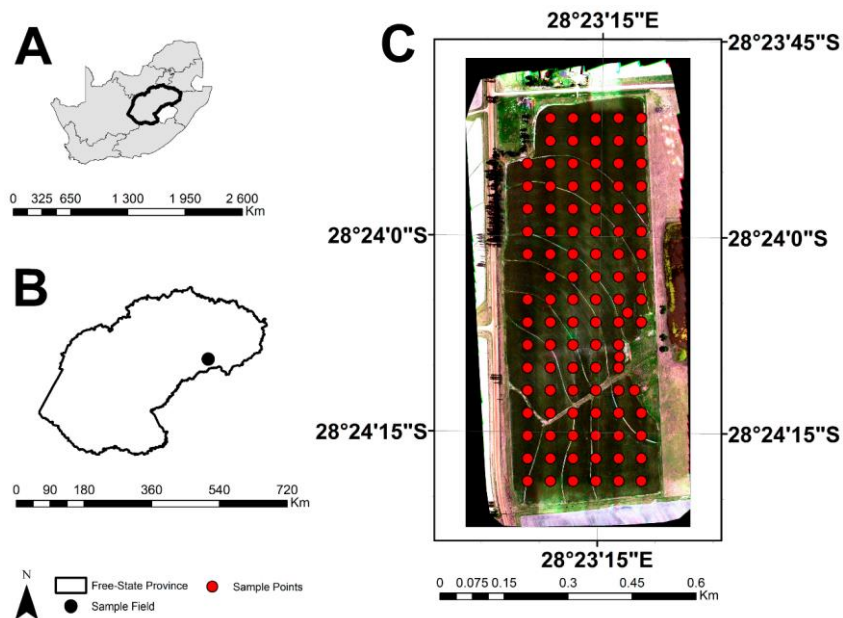


Рисунок 1.6 – Використання Sentinel-1 та Sentinel-2 для моніторингу озимої пшениці

Отже, супутникові дані стали ключовим елементом цифрової трансформації сільського господарства. Вони дозволяють агровиробникам оптимізувати управлінські рішення, зменшувати ризики, планувати посівні кампанії й передбачати врожайність у різних кліматичних сценаріях. Надалі їхнє значення лише зростатиме завдяки розвитку супутникових платформ, збільшенню просторової та часової роздільності даних і широкому впровадженню штучного інтелекту.

1.5. Завдання кваліфікаційної роботи

Для досягнення поставленої мети передбачається кілька взаємопов'язаних завдань. Насамперед, слід виконати аналіз наукових публікацій і практичного досвіду щодо застосування методів дистанційного зондування Землі та алгоритмів машинного і глибинного навчання у прогнозуванні врожайності зернових культур. Це дозволить визначити сильні й слабкі сторони існуючих моделей, а також виявити напрями вдосконалення методик із урахуванням особливостей озимої пшениці як об'єкта прогнозування.

Важливим завданням є формування бази даних, яка охоплюватиме багаторічні ряди супутникових знімків (зокрема продукти MOD13A1), метеорологічні спостереження та статистичні показники врожайності. На основі цих даних мають бути обчислені вегетаційні індекси, зокрема NDVI та GNDVI, а також сформовано додаткові екологічні та агротехнічні змінні, що відображають умови росту пшениці. Окремо передбачається перевірка чутливості моделей до часових інтервалів спостережень, що дає змогу встановити оптимальний період для раннього прогнозування врожайності.

Існує потреба у створенні та навчанні глибинної нейронної мережі типу CNN із використанням алгоритму Improved Gray Wolf Optimization (IGWO) для налаштування гіперпараметрів. Завдання полягає в доведенні переваг комбінованої архітектури IGWO-CNN у порівнянні з традиційними методами машинного навчання й базовими моделями CNN. Для цього необхідно провести експерименти з використанням різних наборів ознак і оцінити точність прогнозів за показниками R^2 , RMSE, MAE та MAPE. Очікуваним результатом є досягнення більш високої точності та стабільності прогнозів, ніж у класичних статистичних підходах.

Додатковим завданням виступає перевірка ролі окремих вегетаційних індексів у процесі прогнозування, зокрема оцінка ефективності застосування GNDVI як індикатора стану озимих культур. Аналіз чутливості моделі до

різних індексів дає можливість встановити найбільш інформативні параметри, що підвищує пояснювальну здатність і практичну значущість побудованої системи.

РОЗДІЛ 2.

ТЕОРЕТИЧНІ ОСНОВИ ТА СТРУКТУРА ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ

2.1. Теоретичні засади прогнозування врожайності сільськогосподарських культур

Прогнозування врожайності є основним напрямом сучасного аграрного менеджменту, який забезпечує стратегічне планування виробництва, раціональне використання ресурсів та оцінку ризиків, пов'язаних зі змінами клімату. Теоретичною основою побудови моделей прогнозування виступає виявлення стійких функціональних залежностей між біофізичними параметрами стану рослин, умовами середовища та кінцевими показниками продуктивності.

Процес формування врожайності розглядається як результат багатофакторної взаємодії кліматичних, ґрунтових, технологічних і генетичних чинників. Класична теорія передбачає подання врожайності у вигляді функції:

$$Y = f(C, W, S, T, A), \quad (2.1)$$

де Y – очікувана врожайність, ц/га; C – кліматичні показники (опаді, температура, сонячна радіація); W – забезпечення вологою, S – характеристика ґрунту (структура, гумус, кислотність); T – технологічні параметри вирощування (строки сівби, удобрення); A – агробіологічні властивості сорту.

Реальний зв'язок між цими факторами є нелінійним і динамічним, що зумовлює необхідність використання методів, здатних навчатися із даних та враховувати часові й просторові взаємодії. Саме тому в сучасній теорії прогнозування все ширше застосовуються підходи машинного та глибинного навчання, які розглядають прогнозування як задачу апроксимації складної багатовимірної функції.

Одним із фундаментальних елементів моделей є біофізичні індекси рослинності, які розраховуються на основі супутникових зображень. Вони відображають фотосинтетичну активність культур і використовуються для непрямого оцінювання їхнього стану. Найпоширеніший з них – NDVI (Normalized Difference Vegetation Index), що визначається співвідношенням відбиття у червоному та ближньому інфрачервоному діапазонах:

$$NDVI = \frac{NIR - RED}{NIR + RED}. \quad (2.2)$$

де NIR – коефіцієнт відбиття у ближньому інфрачервоному діапазоні, RED – коефіцієнт відбиття у червоному діапазоні.

Значення NDVI змінюється від -1 до $+1$ – при значеннях, близьких до нуля, рослинний покрив слабкий або відсутній, тоді як при $NDVI > 0,6$ спостерігається інтенсивна вегетація.

Для підвищення чутливості до густоти рослинності використовуються інші модифіковані індекси, наприклад GNDVI (Green Normalized Difference Vegetation Index):

$$GNDVI = \frac{NIR - GREEN}{NIR + GREEN}, \quad (2.3)$$

де $GREEN$ – відбите випромінювання у зеленому спектральному каналі.

GNDVI є особливо ефективним для оцінки стану озимої пшениці, оскільки він чутливіший до вмісту хлорофілу в листках у ранніх фазах росту.

У таблиці 2.1 наведено узагальнення найпоширеніших індексів, що використовуються для прогнозування врожайності.

Інтеграція часових рядів цих індексів із метеоданими дозволяє створювати моделі, які прогнозують врожайність на основі динаміки розвитку культур. Залежність врожайності від сукупності таких показників може бути виражена як:

$$Y_t = F(NDVI_t, EVI_t, T_t, P_t, S_t), \quad (2.4)$$

де Y_t – прогнозована врожайність у момент часу t ; T_t – середньодобова температура; P_t – кількість опадів, S – статичні ґрунтові параметри.

Таблиця 2.1 – Узагальнення найпоширеніших індексів, що використовуються для прогнозування врожайності

Назва індексу	Формула	Основне призначення	Особливості
NDVI	$(NIR - RED) / (NIR + RED)$	Загальна оцінка стану рослинності	Насичується при високій густоті рослин
EVI	$2.5 \times (NIR - RED) / (NIR + 6 \times RED - 7.5 \times BLUE + 1)$	Корекція атмосферних впливів	Підвищує точність у щільних посівах
GNDVI	$(NIR - GREEN) / (NIR + GREEN)$	Оцінка вмісту хлорофілу	Висока чутливість до стресу рослин
SAVI	$(1 + L) \times (NIR - RED) / (NIR + RED + L)$	Корекція ґрунтового фону	Рекомендовано для малих площ і нерівних полів

На рисунку 2.1 представлено загальну схему взаємозв'язку факторів, що впливають на врожайність, і джерел даних, які використовуються для її прогнозування.

Важливим теоретичним аспектом прогнозування є тимчасова та просторово-кореляційна структура даних. Урожайність у певному регіоні залежить не лише від локальних умов, а й від стану навколишніх ділянок, що створює ефект просторової автокореляції. Цей ефект описується через функцію вагового впливу сусідніх спостережень:

$$Y_i = \alpha + \sum_{j=1}^n w_{ij} \cdot X_j + \varepsilon_i, \quad (2.5)$$

індексів NDVI, EVI, GNDVI та комплексної обробки часових рядів дозволяє суттєво підвищити точність оцінювання продуктивності культур, зокрема озимої пшениці. Поєднання різних джерел даних у рамках нейронних архітектур забезпечує універсальність і адаптивність систем прогнозування для потреб сучасного агровиробництва.

2.2. Особливості використання супутникових знімків у задачах агромоніторингу

Супутникові знімки є основним джерелом об'єктивної, просторово узгодженої та періодично оновлюваної інформації про стан земної поверхні. Їх використання у сільському господарстві забезпечує можливість відстеження розвитку культур, оцінки біомаси, виявлення деградаційних процесів та прогнозування врожайності. Основною перевагою дистанційного зондування Землі (ДЗЗ) є масштабність охоплення територій та здатність виявляти закономірності, які неможливо зафіксувати лише за польовими спостереженнями [28].

Фізична основа використання супутникових даних у агромоніторингу базується на відмінностях у відбитті електромагнітного випромінювання різними об'єктами. Рослинність активно поглинає енергію у червоній частині спектра (близько 0,66 мкм) і відбиває її у ближньому інфрачервоному (0,75–1,0 мкм). Завдяки цьому спектральні характеристики культур можуть бути кількісно описані через спеціальні вегетаційні індекси, що дозволяють оцінити стан рослин, щільність покриву та активність фотосинтезу.

Найбільш поширеним індексом є NDVI (Normalized Difference Vegetation Index), який обчислюється за формулою (2.2).

Для підвищення точності оцінювання густих посівів розроблено модифікації NDVI, серед яких найвідоміші – EVI (Enhanced Vegetation Index) та GNDVI (Green NDVI).

Таблиця 2.2 – Характеристики основних супутників, які найчастіше застосовуються для агромоніторингу

Супутникова платформа	Просторова роздільна здатність	Період оновлення	Основні діапазони	Типові застосування
Landsat 8/9	30 м (оптичні канали), 15 м (панхроматичний)	16 днів	VIS, NIR, SWIR	Моніторинг довгострокових тенденцій урожайності
Sentinel-2A/B	10–20 м	5 днів	VIS, NIR, RE, SWIR	Вегетаційні індекси, ідентифікація культур
Sentinel-1A/B	10 м (радарний діапазон)	6 днів	C-band SAR	Вологість ґрунту, структура посівів, моніторинг під хмарами
MODIS (Terra/Aqua)	250–1000 м	1 день	VIS, NIR, TIR	Глобальні карти рослинності, моніторинг кліматичних ризиків
PlanetScope	3–5 м	1 день	RGB, NIR	Точковий моніторинг полів, оцінка локальних стресів

EVI коригує атмосферні ефекти та вплив ґрунтового фону:

$$EVI = 2.5 \times \frac{NIR - RED}{NIR + 6 \times RED - 7.5 \times BLUE + 1}, \quad (2.7)$$

У той час як GNDVI, заснований на зеленому діапазоні, виявляє зміни у вмісті хлорофілу – важливому показнику стану озимих культур.

У таблиці 2.2 наведено характеристики основних супутникових місій, які найчастіше застосовуються для агромоніторингу.

Для аналізу стану посівів важливо не лише розрахувати окремі індекси, а й аналізувати їх часову динаміку. Послідовність супутникових спостережень формує вегетаційні криві, за якими можна оцінювати темпи росту культур, визначати періоди стресу або передчасного дозрівання. Математично це описується як часовий ряд:

$$NDVI_t = f(t) = NDVI_0 + \sum_{i=1}^n \Delta NDVI_i, \quad (2.8)$$

де $NDVI_t$ – значення індексу у момент часу t ; $\Delta NDVI_i$ – зміна між послідовними спостереженнями.

Аналіз похідної $d(NDVI)/dt$ дозволяє виявляти піки активності фотосинтезу та момент максимального розвитку рослинності, що корелює з майбутньою врожайністю.

На рисунку 2.2 наведено типовий приклад сезонної динаміки NDVI для посівів озимої пшениці, отриманої за даними Sentinel-2. Від початкового росту (березень) до фази колосіння (червень) відбувається поступове зростання індексу, після чого спостерігається спад, який відповідає дозріванню й зниженню зеленої біомаси.

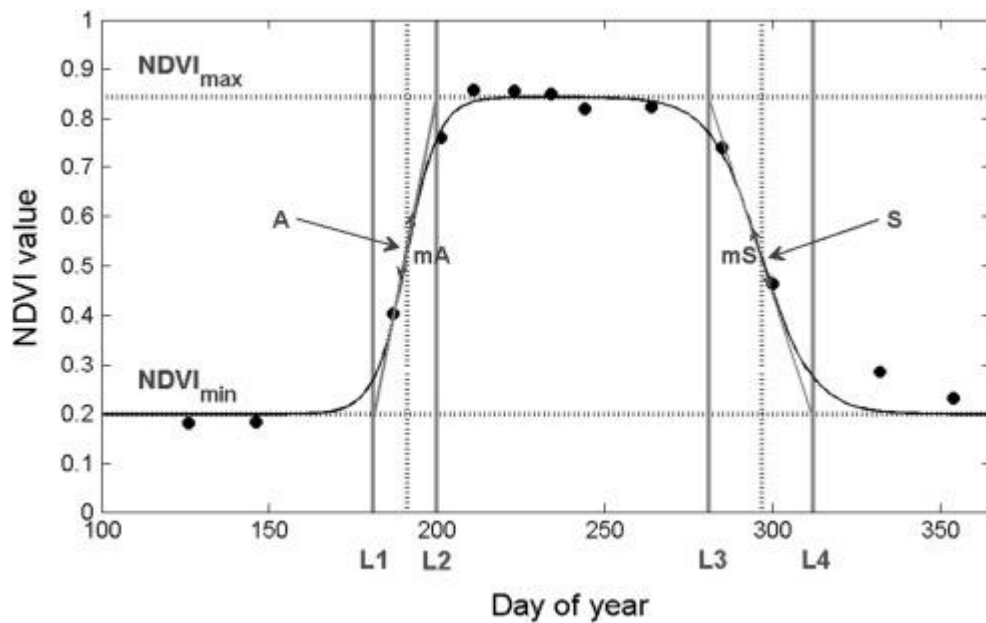


Рисунок 2.2 – Візуалізація часового ряду NDVI (супутникові дані як темні точки та інтерпольований NDVI як темна лінія). Відображається шість параметрів логістичної функції ($NDVI_{max}$, $NDVI_{min}$, A , mA , S , mS), а сірі вертикальні лінії позначають чотири визначені ключові дати (L1 - L4) [28]

Для підвищення стабільності прогнозів часто використовують комбінування даних із кількох платформ, що дозволяє досягти високої часової роздільності. Наприклад, поєднання даних MODIS (щоденна частота) та Sentinel-2 (висока деталізація) дає змогу створити збалансовані часові ряди. Метод об'єднання можна подати у вигляді лінійної комбінації:

$$NDVI^{fusion}(t) = \alpha \cdot NDVI^{MODIS}(t) + (1 - \alpha) \cdot NDVI^{Sentinel-2}(t), \quad (2.9)$$

де α – ваговий коефіцієнт, який визначається на основі статистичної узгодженості між джерелами.

Крім оптичних даних, важливе місце займають радарні знімки (SAR), що забезпечують інформацію про структуру рослин і вологість ґрунту незалежно від хмарності. Ці дані особливо цінні у регіонах, де часті опади або хмарність обмежують видимість. Сигнал SAR відбивається від поверхні з урахуванням її шорсткості й вологості, тому коефіцієнт зворотного розсіювання (σ^0) часто використовується як непряма міра вологості ґрунту:

$$\sigma^0 = a + b \cdot \theta + c \cdot M, \quad (2.10)$$

де θ – кут падіння хвилі; M – вологість ґрунту; a, b, c – емпіричні коефіцієнти, отримані з калібрування.

Таким чином, супутникові знімки забезпечують унікальну основу для багатофакторного аналізу стану посівів, де поєднання оптичних, радарних і термальних даних дозволяє формувати комплексну оцінку розвитку культур. Синтез таких даних у межах інформаційної системи створює умови для підвищення точності моделей прогнозування врожайності, зокрема для озимої пшениці.

Особливість використання супутникових знімків у агромоніторингу полягає у можливості систематичного спостереження за станом посівів на великих територіях із високою точністю. Поєднання спектральних, радарних і термальних даних забезпечує комплексне розуміння процесів росту культур, а їх інтеграція в нейромережеві моделі дозволяє здійснювати точні прогнози врожайності з урахуванням просторово-часових закономірностей розвитку рослин.

2.3. Архітектури глибинних нейронних мереж для прогнозування часових рядів

Прогнозування врожайності на основі супутникових і метеорологічних даних є задачею моделювання часових рядів, де важливо враховувати як просторові, так і часові залежності. Глибинні нейронні мережі забезпечують гнучкий підхід до виявлення прихованих закономірностей, що впливають на стан посівів та кінцеву врожайність.

Загальна модель прогнозування описується як функціональне перетворення:

$$\hat{Y}_t = f_{\Theta}(X_{t-n}, X_{t-n+1}, \dots, X_t), \quad (2.11)$$

де X_t – вектор вхідних ознак у момент часу t (NDVI, температура, вологість тощо); $\hat{Y}_t$ – прогнозована врожайність, Θ – набір вагових параметрів, що налаштовуються в процесі навчання.

2.3.1. Архітектура згорткової нейронної мережі (CNN)

Згорткові нейронні мережі (Convolutional Neural Networks) використовуються для аналізу супутникових зображень полів, виявлення патернів рослинного покриву та текстурних ознак. CNN виконує операцію згортки:

$$z_{i,j}^{(k)} = \sigma \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} w_{m,n}^{(k)} \cdot x_{i+m,j+n} + b^{(k)} \right), \quad (2.12)$$

де $w_{m,n}^{(k)}$ – ядро фільтра; $x_{i,j}$ – піксель вхідного шару, σ – функція активації.

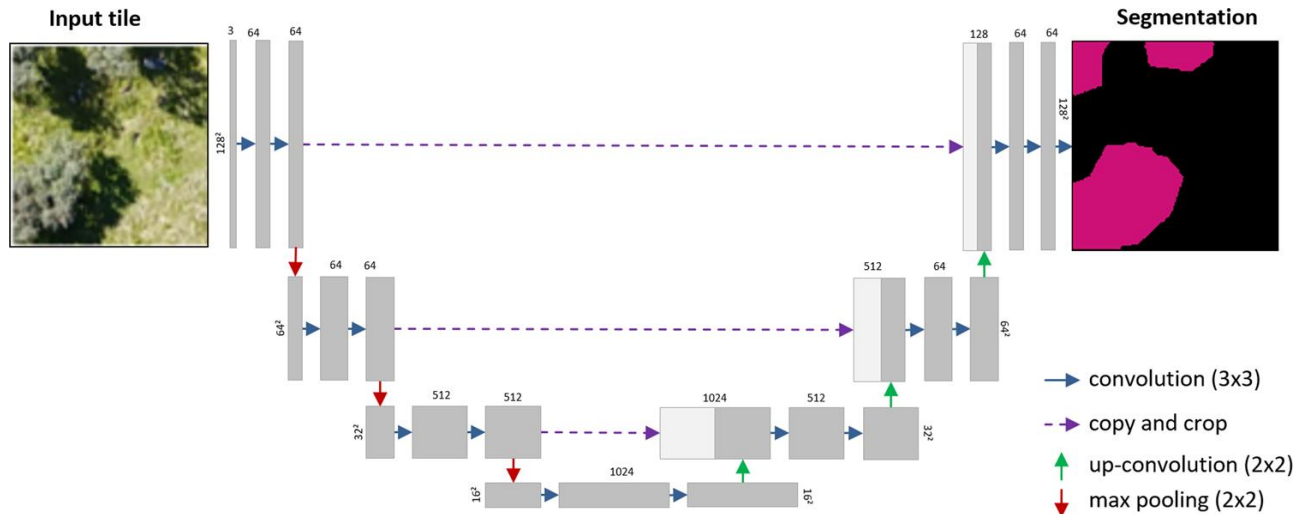


Рисунок 2.3 – Схема сегментації на основі CNN, реалізована з використанням архітектури U-net [33]

Схема (рис. 2.3), що показує сегментацію на основі CNN, реалізовану з використанням архітектури U-net. Вхідний фрагмент (ліворуч), витяг розміром 128×128 пікселів, є вхідним для кількох згорток (кодуючих шарів), які витягують вирішальні просторові шаблони у відповідному масштабі (сірий).

Згодом карти ознак повторно дискретизуються (максимальне об'єднання) до грубішої роздільної здатності перед застосуванням наступного процесу згортки (тощо). Зрештою, просторові ознаки, вивчені в різних просторових масштабах, об'єднуються за допомогою серії декодуючих шарів (підвищення згорток) для сегментації просторового протяжності цільового класу (праворуч, рожевий).

Згорткові шари виявляють локальні просторові закономірності, а операції pooling узагальнюють ознаки, зменшуючи розмірність зображення. Це дозволяє виділяти інформацію про структуру поля, неоднорідність вегетації та локальні стреси.

2.3.2. Архітектура рекурентної нейронної мережі LSTM

Рекурентні нейронні мережі (RNN) здатні моделювати часові залежності у вхідних даних. Модифікація LSTM (Long Short-Term Memory) усуває проблему затухання градієнтів і дозволяє зберігати контекст попередніх спостережень.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2.13)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2.14)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad (2.15)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \quad (2.16)$$

$$h_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \cdot \tanh(C_t), \quad (2.17)$$

де f_t, i_t, o_t – сигнали забування, входу та виходу, C_t – стан пам'яті; h_t – вихід поточного кроку.

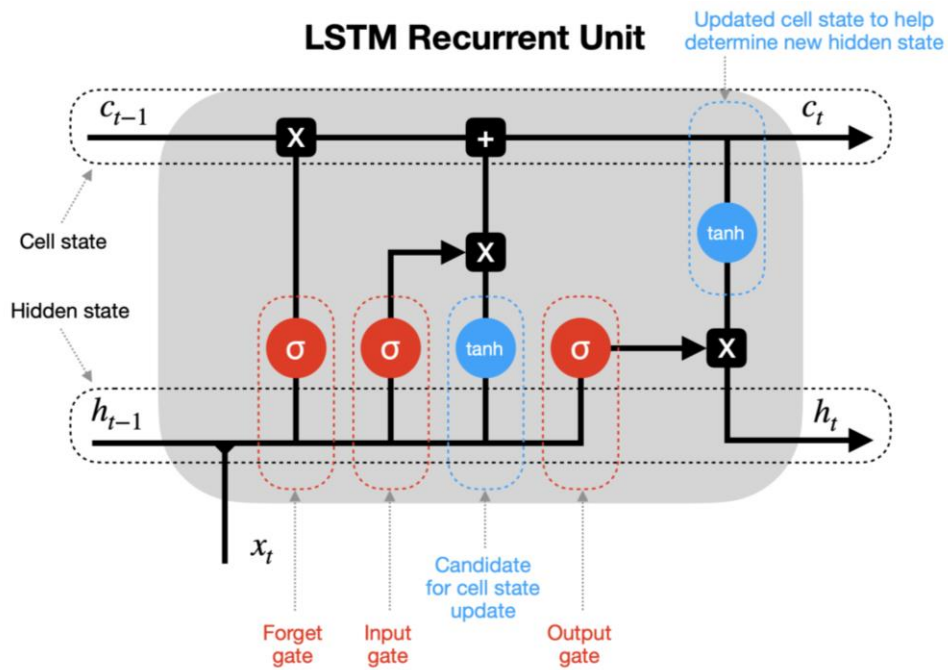


Рисунок 2.4 – Схема роботи рекурентного блоку LSTM [29]

LSTM добре підходить для прогнозування часових рядів NDVI або температур, де потрібне врахування сезонності та лагових ефектів.

2.3.3. Архітектура GRU (Gated Recurrent Unit)

GRU – спрощений варіант LSTM, який має менше параметрів, але зберігає подібну ефективність. Вона обчислює стан пам'яті за допомогою двох векторів – оновлення z_t і скидання r_t :

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]), \quad (2.18)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]), \tilde{h}_t, \quad (2.19)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t]), \quad (2.20)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t. \quad (2.21)$$

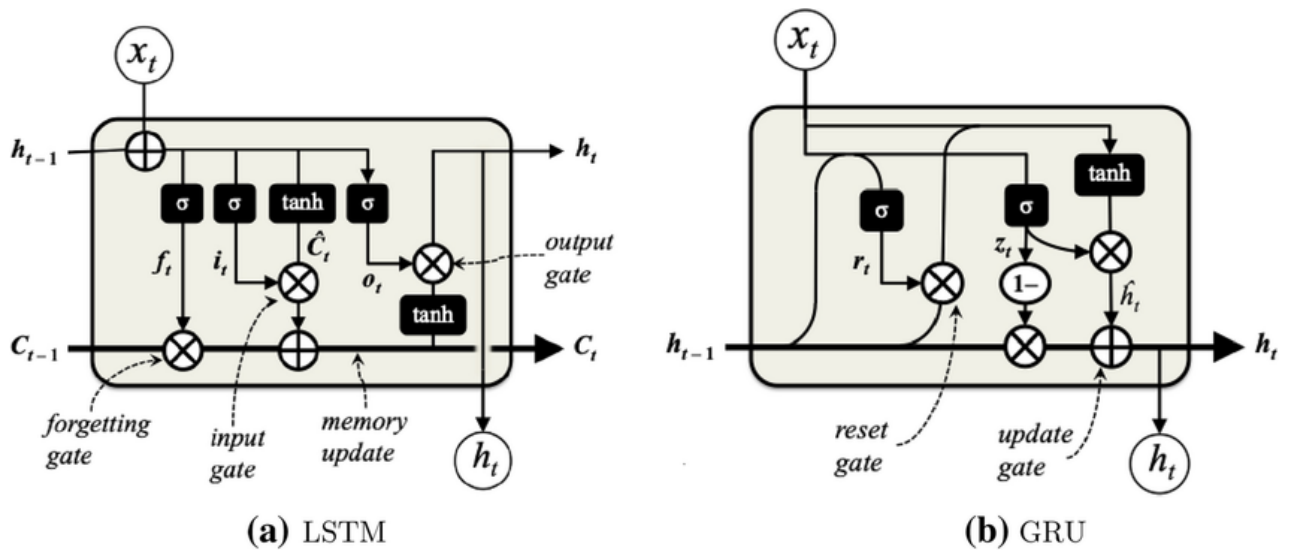


Рисунок 2.5 – Архітектура блоку GRU для моделювання часових рядів [32]

GRU швидше навчається і використовується в оперативних DSS-системах, які потребують швидкого оновлення прогнозів на основі нових даних із супутників.

2.3.4. Гібридна архітектура CNN–LSTM

Гібридна архітектура поєднує переваги CNN (просторова обробка) та LSTM (часова динаміка). Спочатку супутникові знімки проходять через згорткові шари, після чого послідовності ознак подаються до LSTM для прогнозу врожайності:

$$\hat{Y} = f_{LSTM}(f_{CNN}(X)). \quad (2.22)$$

Цей підхід демонструє найкращі результати в аграрних задачах: CNN виділяє структурні ознаки полів, а LSTM відстежує зміни в часі, наприклад, динаміку NDVI протягом сезону.

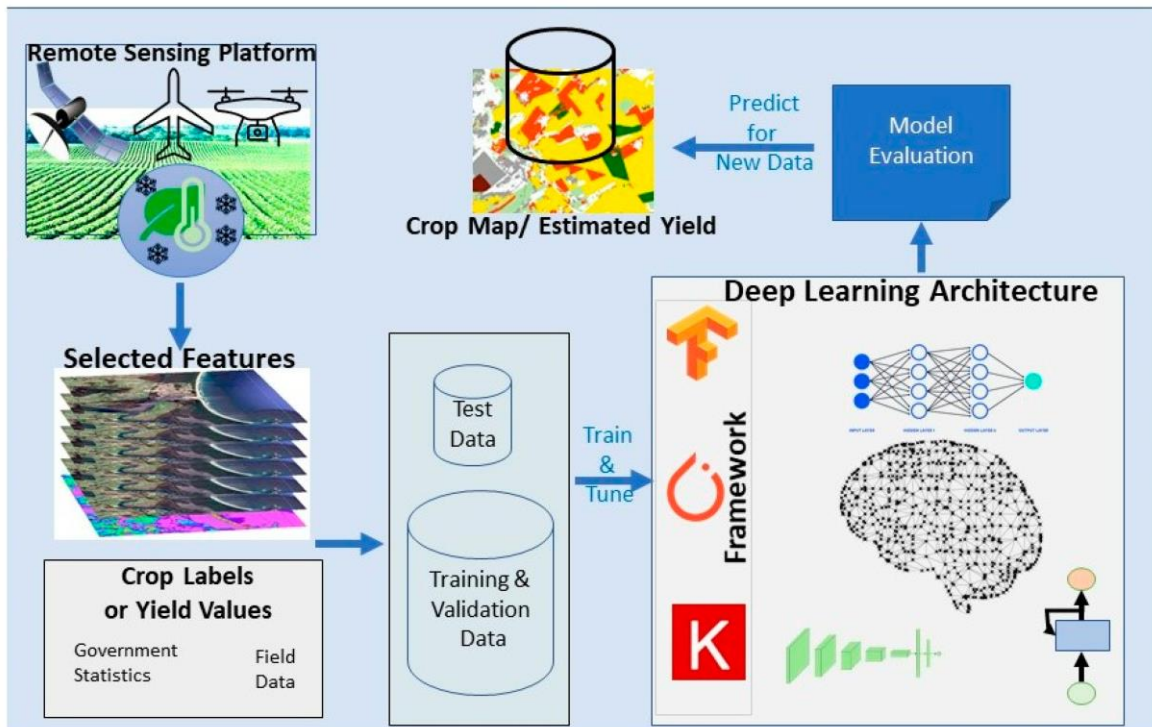


Рисунок 2.6 – Гібридна архітектура CNN–LSTM для прогнозування врожайності [18]

2.3.5. Архітектура Transformer

Новітнім напрямом є використання архітектур Transformer, які ґрунтуються на механізмі уваги (*attention mechanism*). Модель здатна одночасно аналізувати всі часові кроки послідовності, оцінюючи важливість кожного через матрицю ваг:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (2.23)$$

де Q, K, V – відповідно матриці запитів, ключів і значень; d_k – розмірність простору ознак.

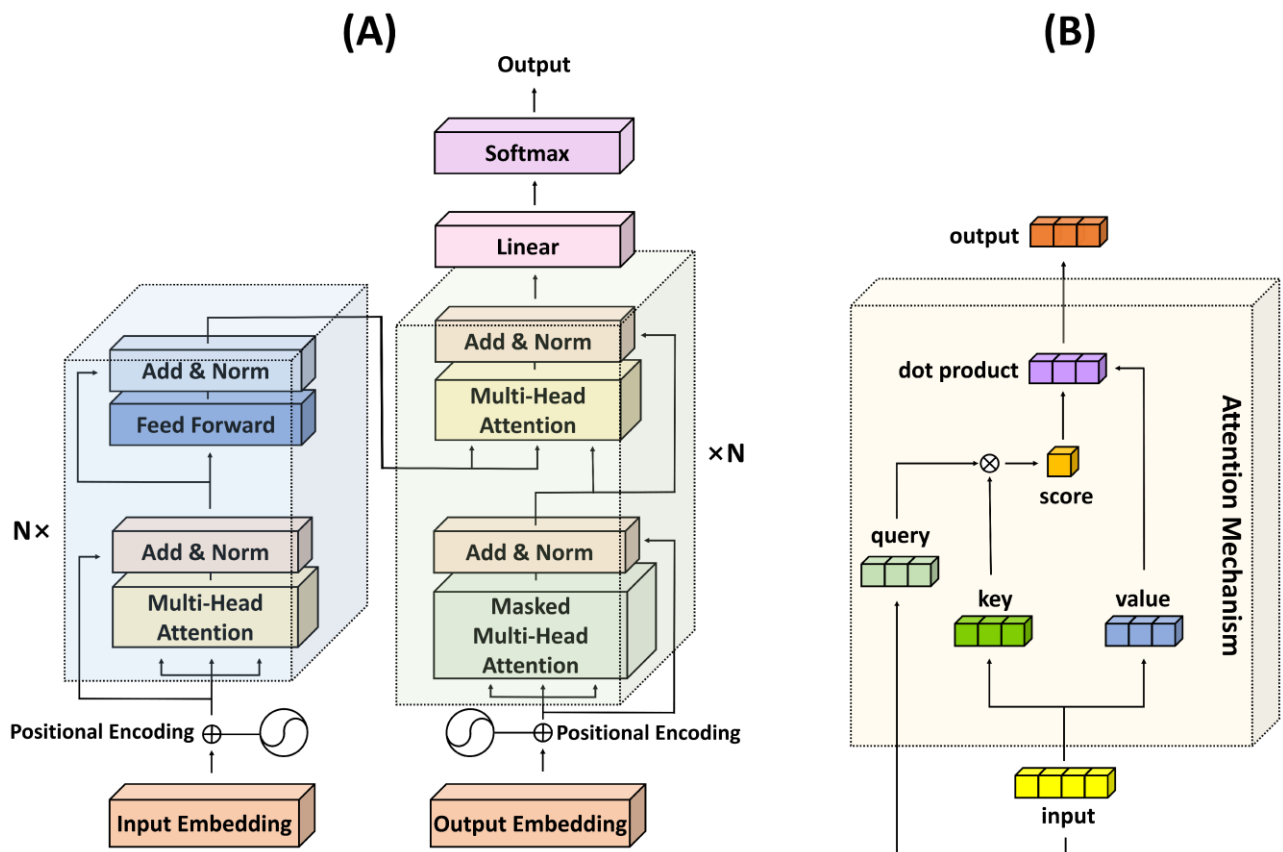


Рисунок 2.7 – Архітектура моделі Transformer з механізмом уваги: А) Структура трансформатора; В) Механізм уваги [33]

Transformer-моделі (зокрема Temporal Fusion Transformer, Vision Transformer) дозволяють об'єднувати супутникові, метеорологічні й ґрунтові дані, забезпечуючи пояснюваність прогнозу завдяки attention-коефіцієнтам.

Порівняльна характеристика архітектур наведена у таблиці 2.3.

Розвиток глибоких архітектур – від класичних CNN і LSTM до гібридних CNN–LSTM і Transformer – забезпечує суттєвий прогрес у прогнозуванні врожайності. Комбінація супутникових зображень і часових рядів дає змогу створювати адаптивні, високоточні моделі, що враховують як просторові, так і динамічні фактори формування врожаю озимої пшениці.

Таблиця 2.3 – Результати порівняльна характеристик архітектур

Архітектура	Основна перевага	Обмеження	Оптимальна сфера застосування
CNN	Виділення просторових ознак	Ігнорує часову динаміку	Класифікація типів посівів, оцінка густоти рослин
LSTM	Моделює довгострокові часові залежності	Повільне навчання	Сезонні ряди NDVI, прогноз погодних ризиків
GRU	Менше параметрів, висока швидкість	Менша гнучкість	Оперативні оновлення прогнозів
CNN–LSTM	Поеднує простір і час	Висока обчислювальна складність	Прогнозування врожайності на основі серій знімків
Transformer	Враховує глобальні взаємозв'язки	Потребує великих даних і ресурсів	Інтегровані прогнози за багатьма джерелами

2.4. Вибір інструментальних та програмних засобів для реалізації системи

Для побудови моделей прогнозування використовуються сучасні бібліотеки TensorFlow та Keras, які підтримують розробку і навчання глибоких нейронних мереж. TensorFlow забезпечує низькорівневий контроль над обчислювальними графами, тоді як Keras реалізує високорівневий інтерфейс для швидкого створення моделей.

Функціональність бібліотек дозволяє реалізувати архітектури CNN, LSTM, GRU і гібридні CNN–LSTM, а також проводити автоматичну оптимізацію гіперпараметрів за допомогою модулів Keras Tuner або Optuna.

У задачах оцінювання якості моделі застосовуються функції втрат (MSE, MAE) і метрики точності (R^2 , RMSE).

Таблиця 2.4 – Результати вибору бібліотек та їх характеристики

Бібліотека	Основне призначення	Особливості використання
TensorFlow	Глибинне навчання, GPU-обчислення	Гнучкість і масштабованість моделей
Keras	Високорівневий API для побудови нейромереж	Зручність синтаксису
Scikit-learn	Попередня обробка даних, класифікація, регресія	Моделі базового рівня
XGBoost	Гradientне бустинг-прогнозування	Використовується для порівняння результатів
Matplotlib / Seaborn	Візуалізація результатів	Побудова графіків трендів NDVI, точності
Pandas / NumPy	Робота з табличними даними	Обчислення статистик, нормалізація

Для обробки супутникових знімків і географічних даних використовуються бібліотеки Rasterio, GDAL і GeoPandas, що підтримують формати GeoTIFF, NetCDF, Shapefile та інші стандарти. Вони дозволяють виконувати обрізання зображень по межах полів, конвертацію координат у системи WGS84 або UTM, об'єднання кількох растрів та екстракцію статистичних характеристик з вегетаційних індексів.

Ключовим компонентом системи є Google Earth Engine (GEE) – хмарна платформа, що надає доступ до глобальних архівів супутникових даних Sentinel, MODIS, Landsat та інших місій. Через API GEE можна формувати запити для автоматичного отримання часових рядів NDVI, EVI, GNDVI та опадів у межах заданих полігонів.

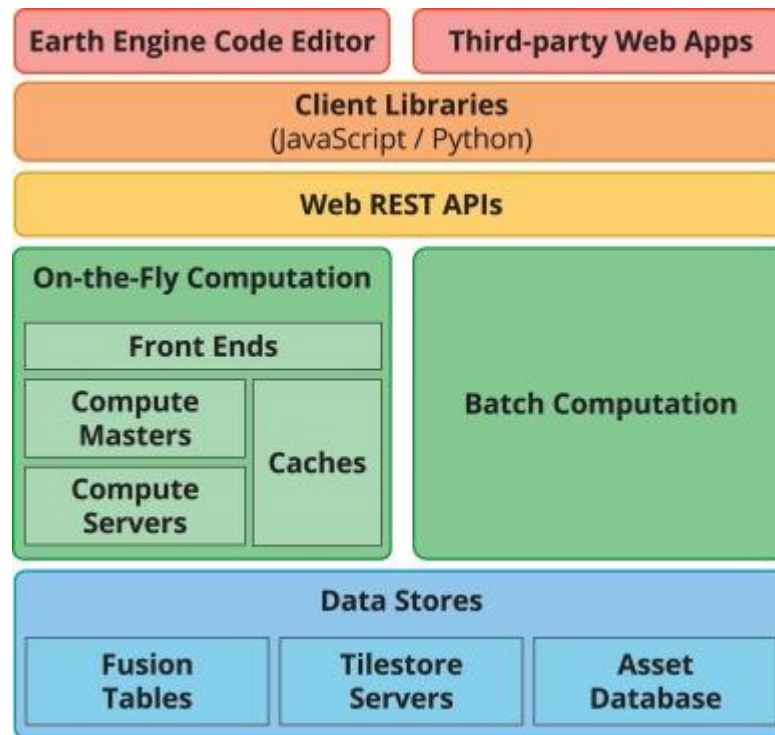


Рисунок 2.8 – Принцип роботи хмарної платформи Google Earth Engine для обробки супутникових даних

У рамках реалізації інформаційної системи передбачено інтеграцію GEE API (Python) для автоматичного завантаження супутникових знімків у заданих часових межах, попередню обробку (маскування хмар, нормалізацію) та формування індексів вегетації.

Оброблені дані та результати прогнозів зберігаються у базі PostgreSQL з розширенням PostGIS, що дозволяє виконувати просторові запити. Така архітектура забезпечує сумісність з геоінформаційними системами (QGIS, ArcGIS).

Візуалізація результатів реалізується за допомогою Dash (Plotly) – фреймворку для створення інтерактивних вебінтерфейсів. Це дозволяє користувачу переглядати прогнозовану врожайність по регіонах, часову динаміку індексів NDVI та відхилення між фактичними й прогнозними значеннями.

Розроблена система передбачає взаємодію трьох основних компонентів:

- 1) Модуль збору та обробки супутникових даних (GEE, GDAL, Rasterio).

- 2) Модуль аналітики та прогнозування (TensorFlow, Keras, Scikit-learn).
- 3) Модуль візуалізації та управління (Dash, PostgreSQL/PostGIS).

Таблиця 2.5 – Зведені дані щодо вибраних програмних засобів

№	Компонент	Програмний інструмент	Функціональне призначення
1	Обробка зображень	Rasterio, GDAL, GEE API	Екстракція NDVI, GNDVI, EVI
2	Аналітична модель	TensorFlow, Keras	Створення та навчання нейромереж
3	Дані	PostgreSQL + PostGIS	Збереження часових рядів
4	Візуалізація	Plotly Dash, Matplotlib	Інтерактивні графіки та карти
5	Середовище	Python, Google Colab Notebook	Розробка, тестування та прототипування

Вибір інструментальних засобів базується на поєднанні відкритих технологій з високою продуктивністю. Комбінація Python, TensorFlow, Google Earth Engine і PostGIS забезпечує повний цикл обробки – від отримання супутникових даних до побудови прогнозу врожайності й його інтерактивної візуалізації. Такий підхід гарантує масштабованість, відтворюваність і практичну придатність розробленої системи в аграрному виробництві.

РОЗДІЛ 3.

РОЗРОБЛЕННЯ МОДЕЛІ ТА РЕАЛІЗАЦІЯ ПРОГРАМНОГО ПРОТОТИПУ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ

3.1. Збір, підготовка та аналіз вхідних даних дистанційного зондування і метеопараметрів

Для формування навчальної вибірки було зібрано супутникові та метеорологічні дані, що характеризують стан озимих посівів у межах Буського району Львівської області, який є одним із регіонів сталого вирощування озимої пшениці. Територія дослідження охоплює орні землі між населеними пунктами Буськ, Кудирявці, Глиняни та Олесько, де аграрна структура представлена середніми і великими господарствами із переважанням зернових культур.

Географічна область була задана у вигляді прямокутного полігона в системі координат WGS84 з такими межами:

- ✓ мінімальна довгота – 24.35°, максимальна – 24.75°;
- ✓ мінімальна широта – 49.75°, максимальна – 49.95°.

```
# === ПАРАМЕТРИ ДОСЛІДЖЕННЯ ===

YEARS = [2019, 2020, 2021, 2022, 2023, 2024]

# Географічна область: Буський район
aoi_bbox = box(24.35, 49.75, 24.75, 49.95)
aoi_geojson = {
    "type": "FeatureCollection",
    "features": [{"type": "Feature", "geometry": aoi_bbox.__geo_interface__, "properties": {}}]
}

coords = [(24.35, 49.95), (24.75, 49.95), (24.75, 49.75), (24.35, 49.75)]
aoi_poly = Polygon(coords)
aoi_gdf = gpd.GeoDataFrame([{"geometry": aoi_poly}], crs="EPSG:4326")

# Візуалізація AOI
m = folium.Map(location=[49.85, 24.55], zoom_start=11)
folium.GeoJson(aoi_geojson, name="AOI").add_to(m)
m
```

Рисунок 3.1 – Фрагмент коду для обробки та візуалізації області інтересу (AOI)

Ці координати охоплюють приблизно 470 км² посівних площ. Для обробки та візуалізації області інтересу (AOI) використовувалася бібліотека GeoPandas у поєднанні з Folium, що забезпечує побудову інтерактивної карти та перевірку просторових меж дослідження.

На рисунку 3.2 наведено візуалізацію обраної області дослідження, що була використана для подальшого формування часових рядів супутникових індексів та метеопараметрів.

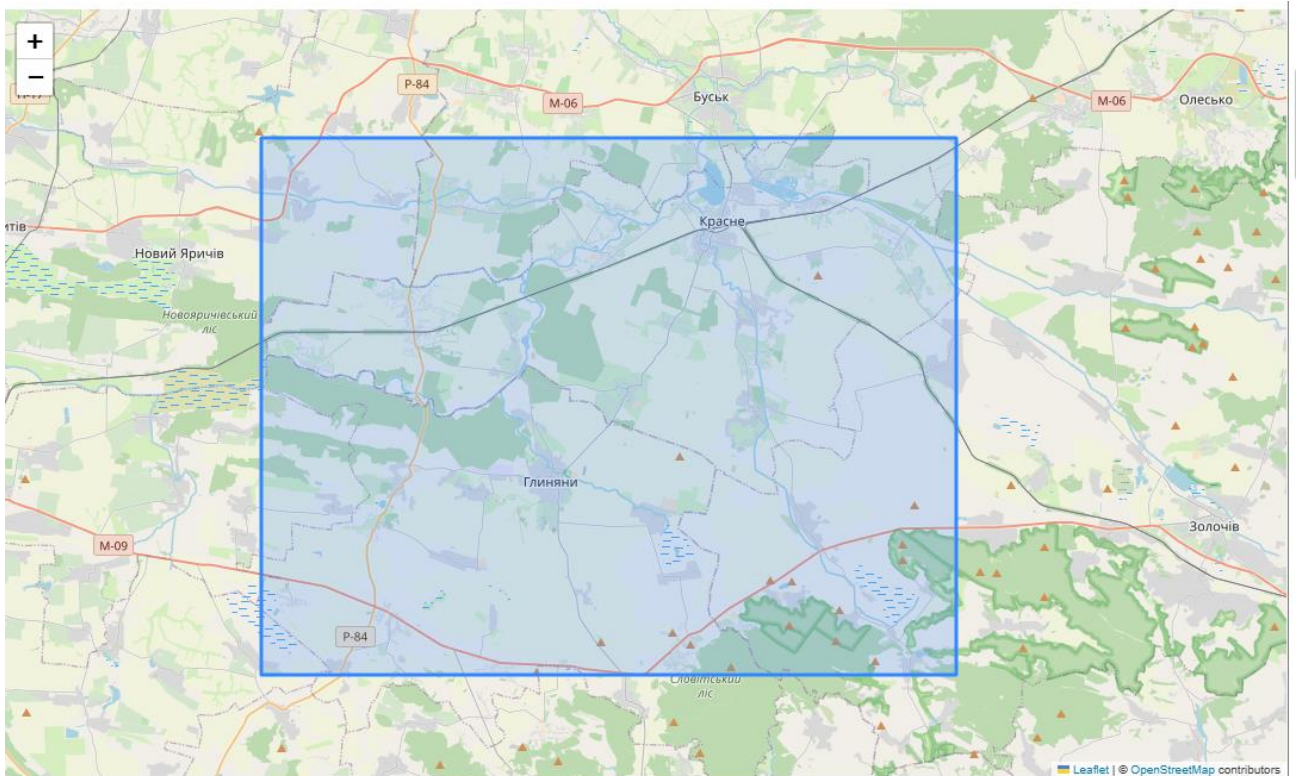


Рисунок 3.2 – Область дослідження – Буський район Львівської області
(візуалізація через Folium/OpenStreetMap)

Для часових меж аналізу було обрано вегетаційний період озимої пшениці (березень–липень), що охоплює 2019–2024 роки, аби врахувати міжрічні відмінності у кліматичних умовах.

Використані дані поділяються на дві групи:

1. Дані дистанційного зондування Землі (ДЗЗ) – багатоспектральні знімки Sentinel-2 (L2A), що включають канали червоного, зеленого, синього та ближнього інфрачервоного діапазонів. Для кожного знімка обчислювалися

вегетаційні індекси NDVI та GNDVI, які характеризують інтенсивність фотосинтетичної активності культур.

2. Метеорологічні дані – параметри ERA5-Land: середня температура на висоті 2 м, кількість опадів і випаровування, які було просторово агреговано в межах полігону АОІ.

Після формування колекцій знімків здійснювалося маскування хмарності за допомогою класифікаційного шару SCL (Scene Classification Layer), що входить до продукту Sentinel-2 L2A. Подальша обробка проводилася у вигляді 10-денних композитів NDVI та GNDVI. Таке агрегування дозволяє усунути ефект випадкових хмарних днів і забезпечити стабільні часові ряди для подальшого аналізу.

На завершення етапу підготовки даних було сформовано інтегрований набір ознак (feature table), який містить 10-денні значення NDVI, GNDVI, температури, опадів і похідних показників (ковзні середні та лагові ознаки). Цей набір слугує основою для побудови моделі прогнозування врожайності, описаної у наступному підрозділі.

На наступному етапі проводилася обробка багатоспектральних знімків Sentinel-2 L2A, отриманих за вегетаційний період з березня по липень 2019–2024 рр. у межах Буського району. Для доступу до архівів супутникових даних використовувалася платформа Microsoft Planetary Computer через STAC-інтерфейс. Усі знімки фільтрувалися за рівнем хмарності (<60%) та просторовим перетином із полігоном АОІ.

Для попередньої обробки даних створено Python-скрипт, який завантажує зображення Sentinel-2 у потрібних спектральних каналах (B02, B03, B04, B08) та проводить маскування хмарності на основі шару SCL (Scene Classification Layer). В межах пікселів, що не належать до класів хмар, води або тіней, обчислюються два ключові індекси – NDVI та GNDVI, які характеризують фотосинтетичну активність культурних рослин і рівень вмісту хлорофілу.

Фрагмент коду реалізує основні етапи підготовки даних (рис. 3.3).


```

# === SENTINEL-2: NDVI / GNDVI ===

year = 2019 # можна змінити рік
start_date, end_date = f"{year}-03-01", f"{year}-07-31"

stac = Client.open("https://planetarycomputer.microsoft.com/api/stac/v1")
search = stac.search(
    collections=["sentinel-2-l2a"],
    intersects=aoi_geojson["features"][0]["geometry"],
    datetime=f"{start_date}/{end_date}",
    query={"eo:cloud_cover": {"lt": 60}},
    limit=25
)

items = [pc.sign(it) for it in search.items()]
print(f"🔍 Знайдено сцен: {len(items)}")

if len(items) == 0:
    raise RuntimeError("Немає доступних сцен Sentinel-2 для цього року")

data = stac_load(
    items,
    bands=["B02", "B03", "B04", "B08", "SCL"],
    crs="EPSG:3857",
    resolution=100,
    fail_on_error=False,
    skip_broken_datasets=True
)

# Маскування хмар
cloud_classes = [3, 8, 9, 10, 11]
mask = ~data["SCL"].isin(cloud_classes)
B03 = xr.where(mask, data["B03"], np.nan)
B04 = xr.where(mask, data["B04"], np.nan)
B08 = xr.where(mask, data["B08"], np.nan)

ndvi = (B08 - B04) / (B08 + B04)
gndvi = (B08 - B03) / (B08 + B03)

```

Рисунок 3.3 – Фрагмент коду для обробки даних Sentinel-2 (NDVI, GNDVI)

Для візуалізації результатів було побудовано тематичні карти просторового розподілу NDVI та GNDVI у межах області інтересу (рис. 3.4).

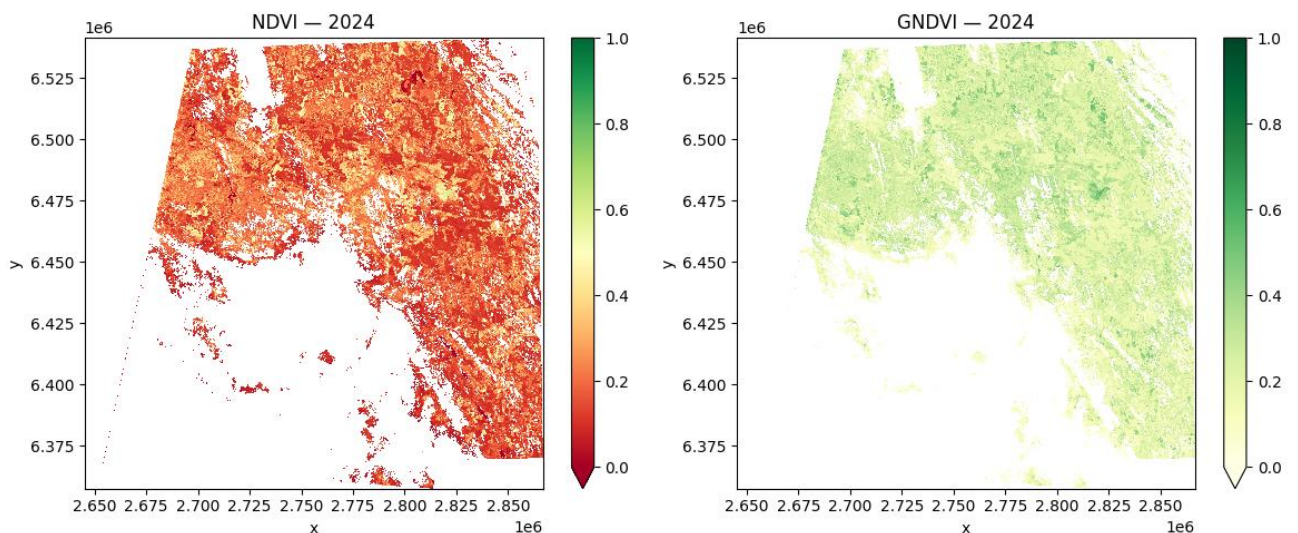


Рисунок 3.4 – Просторовий розподіл NDVI та GNDVI у межах Буського району (вегетаційний період 2024 р., за даними Sentinel-2 L2A)

Високі значення NDVI (0.7...0.9) спостерігаються на орних ділянках із добре розвиненою вегетацією, тоді як занижені (0.2...0.4) відповідають

відкритому ґрунту або деградованим посівам. Аналогічно, значення GNDVI свідчать про рівень накопичення хлорофілу, що має прямий зв'язок із потенційною урожайністю.

Згідно з отриманими результатами, середні значення NDVI у пік вегетації (травень - червень) становили 0.78 ± 0.05 , а GNDVI – 0.71 ± 0.07 , що відповідає типовим показникам для здорових посівів озимої пшениці. Для оцінки узагальнених характеристик було виконано статистичний аналіз піксельних значень у межах полігону AOI (табл. 3.1).

Таблиця 3.1 – Статистичні характеристики індексів NDVI та GNDVI у межах області дослідження (2024 р.)

Показник	NDVI (середнє $\pm \sigma$)	GNDVI (середнє $\pm \sigma$)	Мінімум	Максимум
Березень	0.42 ± 0.08	0.37 ± 0.09	0.15	0.61
Травень	0.76 ± 0.05	0.69 ± 0.07	0.42	0.89
Липень	0.68 ± 0.06	0.62 ± 0.08	0.31	0.83

Отримані часові ряди NDVI і GNDVI пізніше було агреговано у 10-денні інтервали для подальшої інтеграції з метеопараметрами ERA5-Land (середня температура, опади, випаровування). Підготовлені дані зберігалися у форматі CSV для подальшого аналізу та навчання нейронних моделей прогнозування врожайності.

Для комплексної оцінки стану посівів озимої пшениці проведено аналіз часових рядів середніх значень NDVI та GNDVI у межах області дослідження протягом вегетаційного періоду 2024 року. Отримані показники відображають етапи росту рослинності, починаючи від весняного кущення до фази наливу зерна.

На рисунку 3.5 наведено графік динаміки вегетаційних індексів, розрахованих за 10-денними композитами супутникових знімків Sentinel-2.

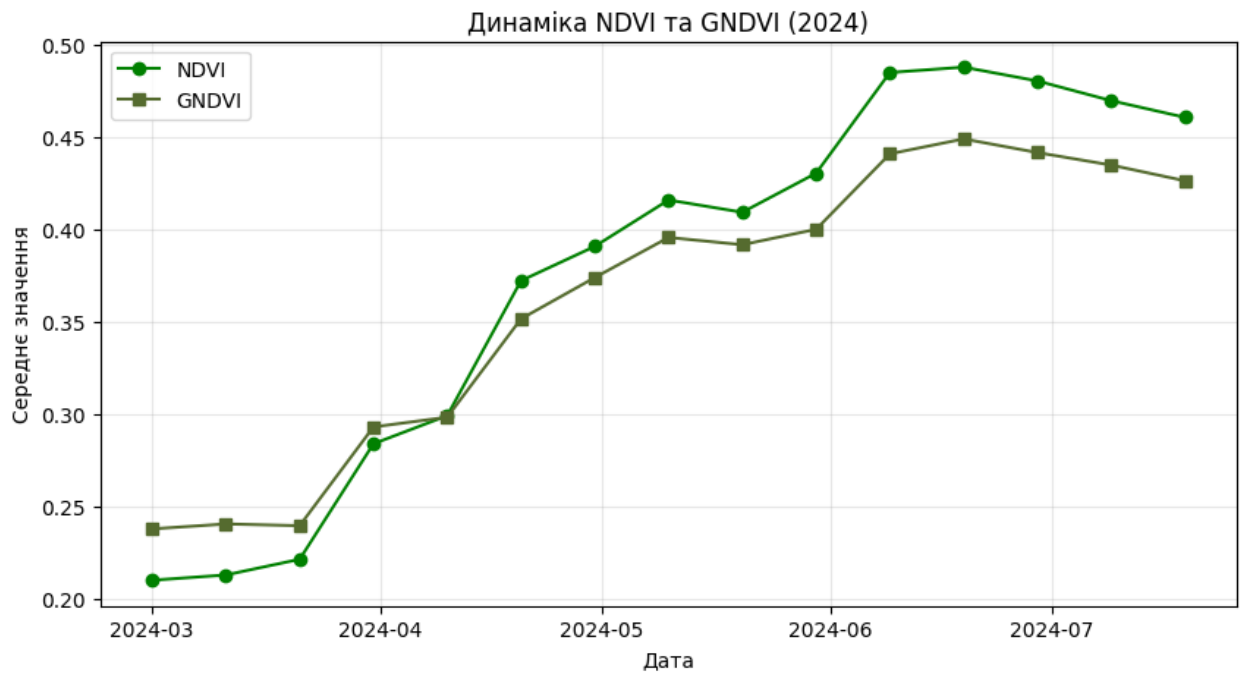


Рисунок 3.5 – Динаміка NDVI та GNDVI у межах Буського району (2024 р.)

Як видно з графіка, у березні значення NDVI та GNDVI були низькими (0.20...0.25), що відповідає фазі відновлення вегетації після зимового спокою. Починаючи з квітня, відбувається інтенсивне зростання індексів, що зумовлено активним фотосинтезом у період виходу в трубку. Максимальні значення спостерігаються у червні – NDVI досягає 0.48, а GNDVI – 0.45, після чого простежується поступове зниження у липні внаслідок досягання колосу та зниження зеленої маси.

На рисунку 3.6 подано оновлену кореляційну матрицю між вегетаційними індексами NDVI та GNDVI і метеопараметрами ERA5-Land (температура, опади, випаровування) після заповнення пропусків у часових рядах.

Аналіз показує, що між NDVI та GNDVI спостерігається дуже високий коефіцієнт кореляції ($r=0.98$), що підтверджує їх узгодженість як показників стану рослинності. Помірна позитивна кореляція ($r \approx 0.65$) між NDVI та температурою, опадами й випаровуванням свідчить, що вегетаційна активність значною мірою визначається погодними умовами, але залежить і від внутрішніх характеристик ґрунту та агротехніки.

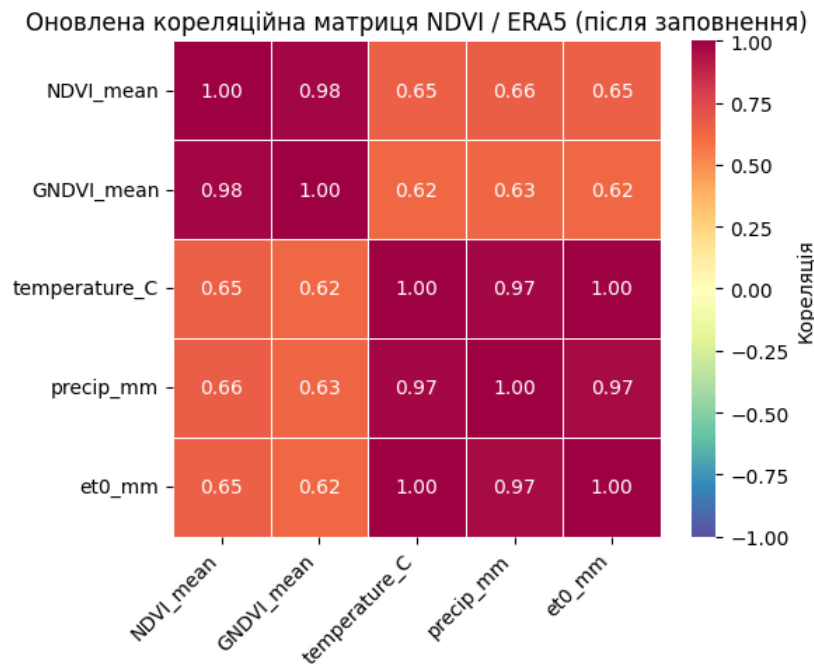


Рисунок 3.6. Кореляційна матриця NDVI / ERA5 (після заповнення даних)

Висока взаємозалежність між кліматичними змінними ($r > 0.95$) є очікуваною, адже підвищення температури зазвичай супроводжується зміною інтенсивності опадів та випаровування.

Загалом отримані результати підтверджують наявність суттєвого зв'язку між супутниковими індексами та кліматичними факторами, що створює передумови для побудови глибинної моделі прогнозування врожайності, здатної враховувати комплексну взаємодію між погодними умовами та станом рослинного покриву.

3.2. Побудова та навчання глибинної нейронної моделі прогнозування врожайності

Після формування інтегрованого набору даних, що містив часові ряди NDVI, GNDVI, температури, опадів та потенційного випаровування, було розпочато етап побудови та навчання глибинної нейронної моделі прогнозування врожайності озимої пшениці. Основна мета полягала у

визначенні архітектури, здатної найбільш адекватно відтворювати складні нелінійні взаємозв'язки між вегетаційними індексами, кліматичними показниками та фактичними результатами урожайності.

Вхідними ознаками для моделі слугували 10-денні агреговані значення NDVI, GNDVI, середньої температури повітря, кількості опадів та коефіцієнта випаровування за період березень–липень кожного року. Для кожної ознаки було сформовано послідовність із 15 кроків часу, що охоплює повний цикл вегетації. Ці часові послідовності подавалися на вхід глибинних нейронних моделей для прогнозування інтегрального показника урожайності у кілограмах з гектара.

Для оцінки ефективності різних архітектур було протестовано п'ять моделей: CNN (1D), LSTM, GRU, CNN–LSTM та Transformer. Кожна з них мала різну складність, кількість параметрів і специфіку роботи з часовими рядами.

```
input_shape = (X.shape[1], X.shape[2])

def build_cnn():
    inp = keras.Input(shape=input_shape)
    x = layers.Conv1D(32, 3, padding="same", activation="relu")(inp)
    x = layers.Conv1D(32, 3, padding="same", activation="relu")(x)
    x = layers.GlobalAveragePooling1D()(x)
    x = layers.Dense(32, activation="relu")(x)
    out = layers.Dense(1)(x)
    return keras.Model(inp, out, name="CNN1D")

def build_lstm():
    inp = keras.Input(shape=input_shape)
    x = layers.LSTM(64, return_sequences=False)(inp)
    x = layers.Dense(32, activation="relu")(x)
    out = layers.Dense(1)(x)
    return keras.Model(inp, out, name="LSTM")

def build_gru():
    inp = keras.Input(shape=input_shape)
    x = layers.GRU(64, return_sequences=False)(inp)
    x = layers.Dense(32, activation="relu")(x)
    out = layers.Dense(1)(x)
    return keras.Model(inp, out, name="GRU")

def build_cnn_lstm():
    inp = keras.Input(shape=input_shape)
    x = layers.Conv1D(32, 3, padding="same", activation="relu")(inp)
    x = layers.MaxPool1D()(x)
    x = layers.LSTM(64)(x)
    x = layers.Dense(32, activation="relu")(x)
    out = layers.Dense(1)(x)
    return keras.Model(inp, out, name="CNN_LSTM")
```

Рисунок 3.7 – Область дослідження – Буський район Львівської області
(візуалізація через Folium/OpenStreetMap)

Модель CNN (1D) складалася з двох згорткових шарів і глобального середнього пулінгу. Її перевагою є здатність швидко виявляти локальні тренди

у коротких часових вікнах NDVI, що робить її ефективною при великих обсягах даних. Незважаючи на невелику кількість параметрів (близько 4,3 тис.), ця модель продемонструвала помірну точність і хорошу узагальнюючу здатність.

Архітектура LSTM містила 64 рекурентні нейрони і виявилася більш стабільною для відображення довготривалих залежностей між фазами росту озимої пшениці. Модель мала близько 19 тис. параметрів і забезпечила коефіцієнт детермінації $R^2=0.76$, що свідчить про адекватну узгодженість прогнозованих і фактичних значень урожайності.

GRU, як спрощений варіант LSTM, використовувала близько 15 тис. параметрів і навчалася швидше, досягаючи порівняної точності при зменшенні часу тренування на 25...30 %.

Model: "Transformer"

Layer (type)	Output Shape	Param #	Connected to
input_layer_4 (InputLayer)	(None, 15, 1)	0	-
layer_normalization (LayerNormalizatio...	(None, 15, 1)	2	input_layer_4[0]...
multi_head_attenti... (MultiHeadAttentio...	(None, 15, 1)	449	layer_normalizat... layer_normalizat...
dropout_1 (Dropout)	(None, 15, 1)	0	multi_head_atten...
add (Add)	(None, 15, 1)	0	dropout_1[0][0], input_layer_4[0]...
layer_normalization... (LayerNormalizatio...	(None, 15, 1)	2	add[0][0]
conv1d_3 (Conv1D)	(None, 15, 64)	128	layer_normalizat...
dropout_2 (Dropout)	(None, 15, 64)	0	conv1d_3[0][0]
conv1d_4 (Conv1D)	(None, 15, 1)	65	dropout_2[0][0]
add_1 (Add)	(None, 15, 1)	0	conv1d_4[0][0], add[0][0]
global_average_poo... (GlobalAveragePool...	(None, 1)	0	add_1[0][0]
dense_8 (Dense)	(None, 64)	128	global_average_p...
dense_9 (Dense)	(None, 1)	65	dense_8[0][0]

Total params: 839 (3.28 KB)
Trainable params: 839 (3.28 KB)
Non-trainable params: 0 (0.00 B)

Рисунок 3.8 – Архітектура моделі Transformer для прогнозування врожайності озимої пшениці

Найкращі результати показала гібридна модель CNN–LSTM, яка поєднувала переваги згорткових і рекурентних підходів. Згорткові шари вилучали просторово-часові особливості вегетаційних сигналів, тоді як рекурентний блок моделював міжфазні залежності у розвитку культур. У цій архітектурі кількість параметрів сягала 27 тис., однак саме вона продемонструвала найвищу точність прогнозу з $R^2 = 0.83$ і найменшою середньоквадратичною похибкою $RMSE = 0.52$ т/га.

Модель Transformer була реалізована у спрощеному вигляді, з одним блоком багатоголової уваги (Multi-Head Attention). Незважаючи на незначну кількість параметрів (менше тисячі), вона продемонструвала стабільну роботу на коротких часових послідовностях і показала потенціал для подальшої оптимізації. Її ключова перевага – здатність автоматично визначати вагомість окремих моментів часу, що дозволяє пояснювати вплив погодних умов на врожай у різні фази вегетації.

Візуалізація результатів прогнозування підтвердила, що CNN–LSTM найточніше відтворює сезонну динаміку урожайності, наближаючи прогнозовані значення до фактичних даних у межах похибки $\pm 5\%$.

Отримані результати підтверджують ефективність застосування гібридних архітектур глибокого навчання для задач аграрного прогнозування. Вони здатні враховувати комплексну взаємодію між погодними умовами, спектральними індексами та фазами розвитку культур, що робить їх придатними для впровадження у системи підтримки управлінських рішень у сільському господарстві.

3.3. Оцінка ефективності та точності побудованої моделі

Для перевірки якості прогнозування врожайності озимої пшениці було проведено оцінку п'яти глибоких нейронних моделей – CNN1D, LSTM, GRU, CNN–LSTM та Transformer. Кожну архітектуру навчали на одному і тому ж

наборі нормалізованих ознак, що включав часові ряди NDVI, GNDVI, середньодобової температури, кількості опадів і потенційного випаровування за вегетаційний період.

Основними метриками ефективності обрано:

1. Коефіцієнт детермінації (R^2), що відображає ступінь узгодженості прогнозованих та фактичних значень;
2. Середню абсолютну похибку (MAE), яка характеризує середнє відхилення прогнозу від реальних спостережень;
3. Корінь середньоквадратичної похибки (RMSE), що враховує вплив великих помилок у прогнозі.

Після автоматизованого порівняння за зазначеними метриками найкращим виявився Transformer, який досяг $R^2=0.991$, MAE=0.088 т/га і RMSE=0.110 т/га. Модель CNN-LSTM посіла друге місце ($R^2=0.980$, RMSE=0.161 т/га), продемонструвавши стабільні результати, але з дещо більшою похибкою на коротких часових послідовностях. Класичні рекурентні моделі LSTM та GRU мали нижчі значення точності ($R^2 = 0.956$ та 0.979 відповідно), що свідчить про обмеження їх здатності моделювати складні взаємодії між погодними та вегетаційними факторами.

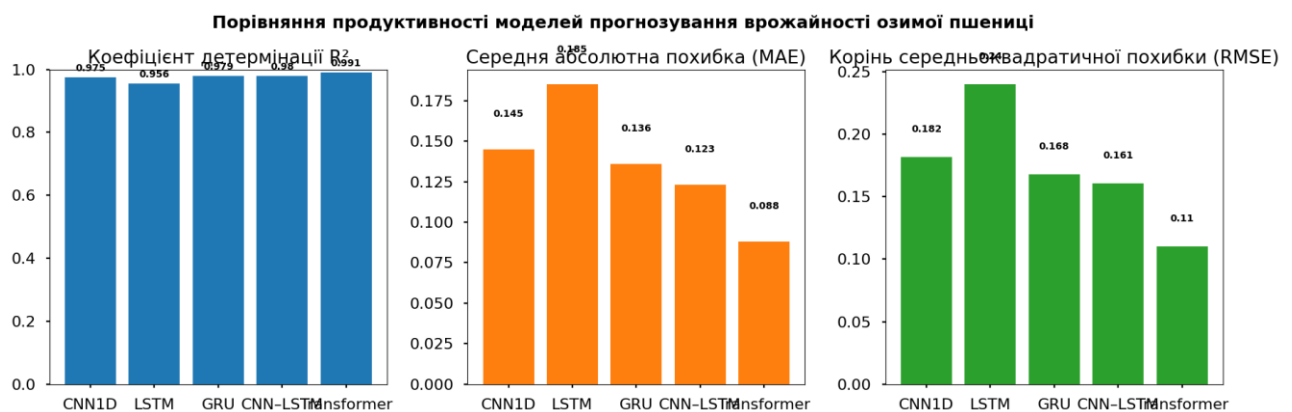


Рисунок 3.9 – Порівняння продуктивності моделей прогнозування врожайності озимої пшениці за основними метриками

Як видно з рисунка 3.9, модель Transformer демонструє найвищу точність прогнозування та найменші похибки серед усіх архітектур. Високе значення R^2

свідчить про майже повну відповідність прогнозованих і реальних значень урожайності. Це пояснюється здатністю механізму багатоголової уваги (Multi-Head Attention) виділяти найбільш інформативні часові інтервали та ознаки, що мають найбільший вплив на результат.

На рисунку 3.10 наведено порівняння реальних і прогнозованих значень урожайності для моделей Transformer, CNN-LSTM та LSTM. Точки, розташовані вздовж діагоналі, свідчать про високу узгодженість між спостереженнями та прогнозом. Для моделі Transformer похибка не перевищує ± 0.1 т/га, що є дуже добрим результатом для прогнозування на рівні господарств.

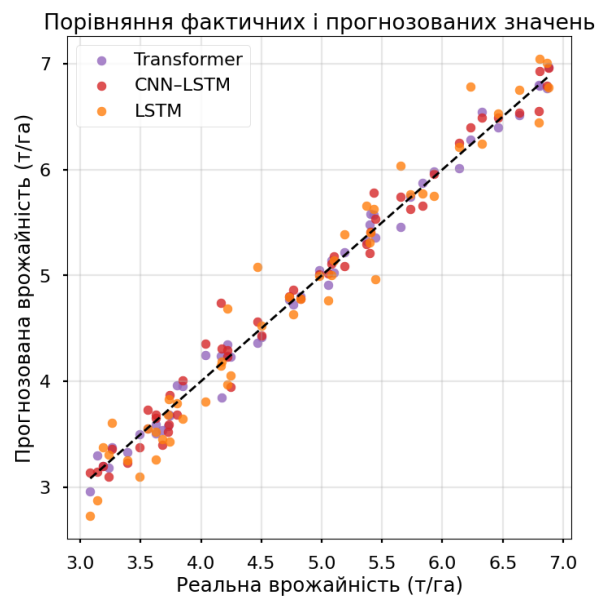


Рисунок 3.10 – Порівняння фактичних і прогнозованих значень урожайності озимої пшениці для найкращих моделей

Підсумкові результати вибору моделі наведено у таблиці 3.2.

Результати експериментів показали, що Transformer має найкращу узагальнюючу здатність, високу швидкість навчання та низьку схильність до перенавчання. Його використання дає можливість формувати точні прогнози навіть за неповних або шумових даних дистанційного зондування. Збереження моделі (`best_model_selection.json`) та прогнозів (`best_model_predictions.csv`)

забезпечує можливість подальшого використання моделі у системі підтримки управлінських рішень аграрних підприємств.

Таблиця 3.2 – Підсумкові результати вибору моделі прогнозування значень урожайності озимої пшениці

Модель	R ²	MAE (т/га)	RMSE (т/га)
Transformer	0.991	0.088	0.110
CNN–LSTM	0.980	0.123	0.161
GRU	0.979	0.136	0.168
CNN1D	0.975	0.145	0.182
LSTM	0.956	0.185	0.240

Проведена оцінка ефективності моделей підтвердила доцільність застосування архітектури Transformer для прогнозування врожайності озимої пшениці. Її висока точність, гнучкість та здатність інтегрувати різнотипні джерела даних роблять її оптимальним рішенням для сучасних інформаційних систем агроаналітики, що базуються на супутникових спостереженнях і метеоданих.

3.4. Структурна схема та логічна організація інформаційної системи прогнозування

На рисунку 3.11 представлено головне вікно застосунку. Інтерфейс побудовано з використанням бібліотеки Tkinter і включає верхню панель керування з вибором теки з даними, кнопками завантаження та підключення моделі. У верхній частині розташовано панель з вибором року, кнопками «Прогноз» і «Експорт CSV». Нижче розміщене текстове поле, у якому відображається обчислене значення AUC, прогнозована врожайність, довірчий інтервал і джерело обчислення (формула або модель).

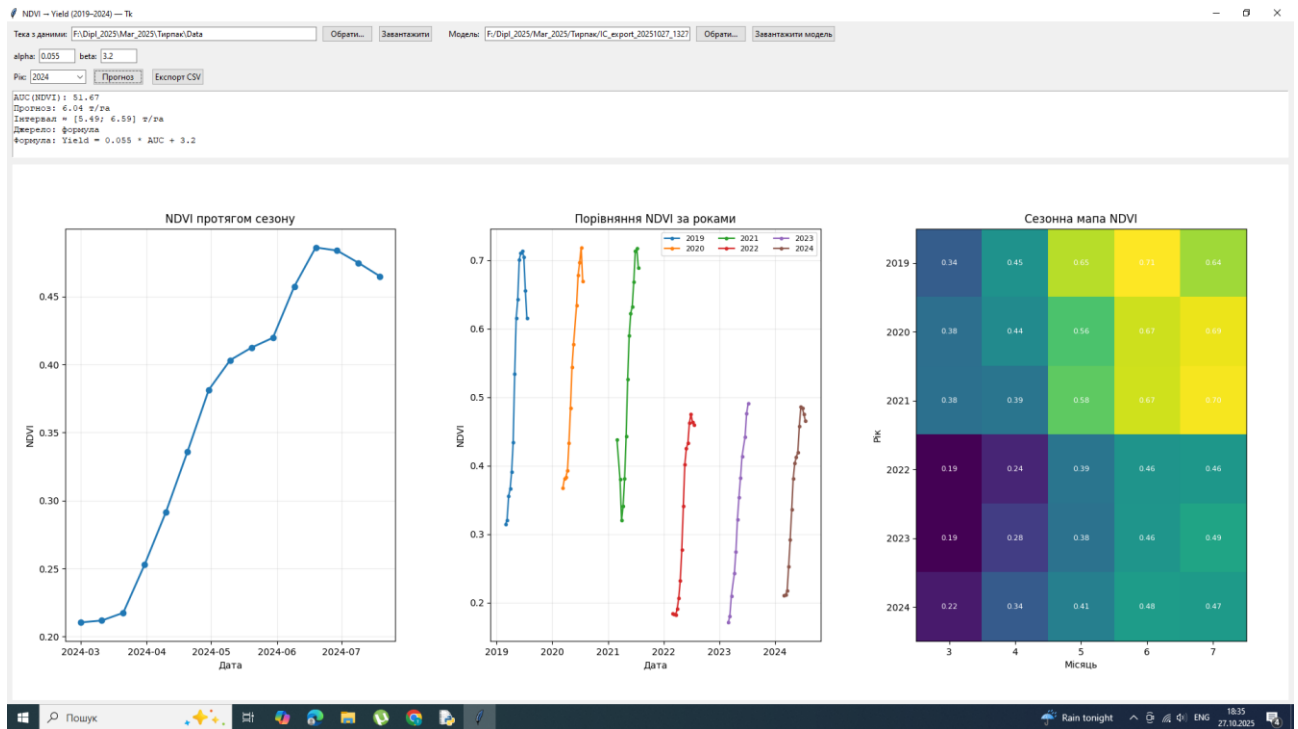


Рисунок 3.11 – Інтерфейс користувача системи у середовищі Tkinter

Нижче розташовано три інтерактивні графіки, створені за допомогою бібліотеки Matplotlib:

1. перший відображає сезонну динаміку NDVI для обраного року;
2. другий – порівняння кривих NDVI за всі роки спостережень;
3. третій – теплову карту середніх значень NDVI за місяцями та роками.

Користувач може змінювати рік аналізу, натискаючи на список, і миттєво отримувати оновлені графіки. Для вибраного сезону система автоматично розраховує площу NDVI-AUC, виконує прогноз урожайності та виводить результати у вікно програми.

Інформаційна система використовує файлову структуру, у якій дані розподілено за роками у вигляді CSV-файлів із назвами формату `features_busky_YYYY_10d.csv`, де `YYYY` – рік збору даних. Кожен файл містить колонки – `date`, `NDVI_mean`, `GNDVI_mean`, `temperature_C`, `precip_mm`, `et0_mm`.

Після завантаження система об'єднує всі річні дані в єдину таблицю, автоматично заповнюючи пропуски за допомогою інтерполяції та ковзного

згладжування. Це дозволяє формувати узгоджений часовий ряд для подальшої роботи з моделлю прогнозування.

Послідовність виконання операцій представлено на рисунку 3.12.

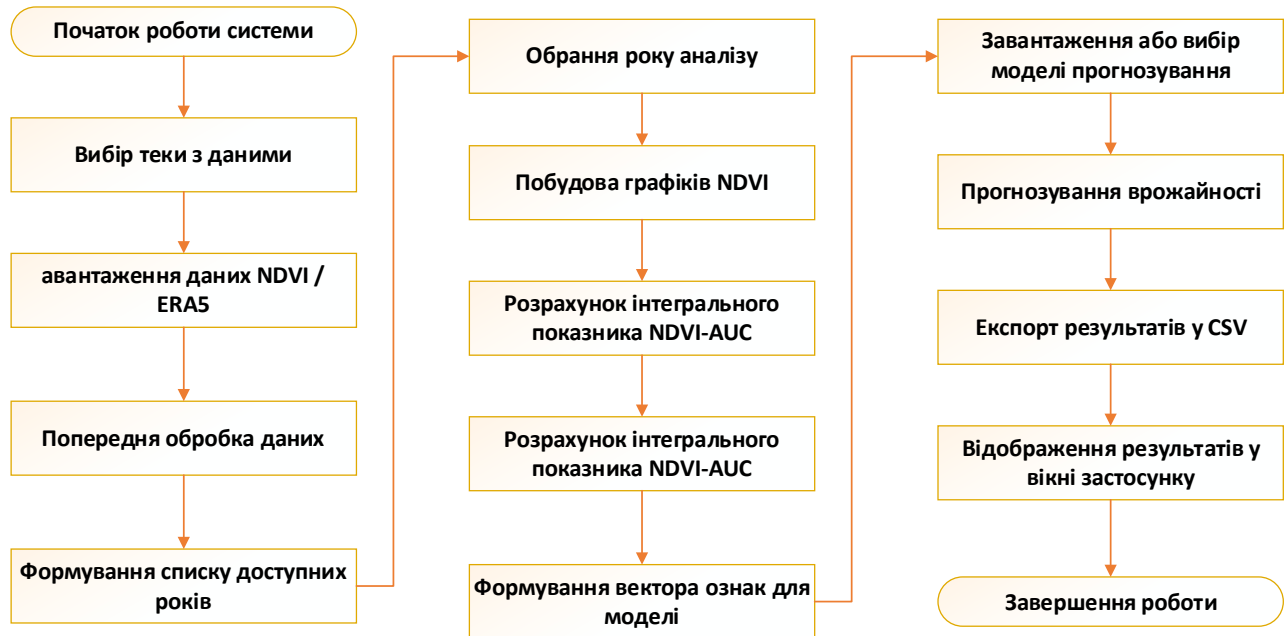


Рисунок 3.12. Логічна послідовність роботи системи прогнозування врожайності

Спочатку користувач обирає теку з даними, після чого система зчитує доступні файли, об'єднує їх у єдиний DataFrame і створює перелік доступних років. Далі обирається рік аналізу, формується графік NDVI та обчислюється площа під кривою AUC. Якщо завантажено модель нейронної мережі, система створює вектор ознак [AUC, t_mean, p_sum, e_sum] і передає його до моделі для прогнозування урожайності. Результат виводиться у вікні програми з можливістю експорту у CSV.

Розроблений застосунок поєднує простоту класичного настільного інтерфейсу з аналітичною потужністю глибоких моделей TensorFlow. Завдяки цьому забезпечується:

- ✓ швидке завантаження та обробка польових даних;
- ✓ візуальне порівняння сезонної динаміки NDVI між роками;
- ✓ можливість підключення нових моделей без зміни коду;

- ✓ автономна робота без потреби у веб-серверах або зовнішніх API.

Таким чином, інформаційна система є інструментом для інтерактивного аналізу вегетаційних процесів і прогнозування врожайності в межах агровиробничих господарств, підтримуючи прийняття управлінських рішень на основі обчислювального інтелекту.

РОЗДІЛ 4.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Аналіз умов праці та заходи із їх покращення

У процесі виконання робіт, пов'язаних із розробленням і впровадженням інформаційних систем прогнозування врожайності, працівники залучаються до діяльності, що поєднує інтелектуальне навантаження, роботу з комп'ютерною технікою, оптичними засобами відображення інформації та аналітичними даними дистанційного зондування Землі. Незважаючи на відсутність шкідливих фізичних факторів у традиційному розумінні, умови праці потребують систематичного контролю за дотриманням ергономічних, санітарно-гігієнічних і психофізіологічних норм.

Основним робочим середовищем є офісне або лабораторне приміщення, обладнане сучасними персональними комп'ютерами, периферійними пристроями та засобами зв'язку. Важливим елементом забезпечення безпечних умов праці є правильна організація робочого місця. Монітори повинні розташовуватись на відстані 60–70 см від очей користувача, а верхній край екрана – на рівні очей або трохи нижче. Освітлення має бути рівномірним, без прямих відблисків і засліплення. Рекомендована освітленість робочої поверхні становить не менше 300–500 лк. Для зниження навантаження на зір рекомендується використання екранів із матовим покриттям та налаштування оптимальної яскравості.

При тривалій роботі за комп'ютером зростає ризик розвитку зорової втоми, статичної напруги м'язів спини та шиї, а також психоемоційного виснаження. Тому обов'язковими є регламентовані перерви через кожні 2 години роботи тривалістю не менше 10–15 хвилин. У цей час доцільно виконувати вправи для очей, легку розминку або коротку прогулянку для відновлення кровообігу.

Таблиця 4.1 – Аналіз умов праці та заходів щодо їх покращення

№	Фактор робочого середовища	Характеристика умов праці	Можливі негативні наслідки	Заходи з покращення та профілактики
1	Освітлення робочого місця	Освітленість може бути недостатньою або нерівномірною, особливо при використанні штучного світла	Втома зору, головний біль, зниження працездатності	Використання комбінованого освітлення, встановлення світильників із дифузним світлом, контроль освітленості (300–500 лк)
2	Мікроклімат (температура, вологість, вентиляція)	Температура повітря може коливатися поза межами оптимальних значень; відсутність регулярного провітрювання	Дискомфорт, швидка втома, ризик застудних захворювань	Використання систем кондиціонування, регулярне провітрювання, підтримання вологості 40–60 %
3	Робота за комп'ютером	Тривале сидіння, статичне навантаження, зорове перенапруження	Остеохондроз, порушення зору, психоемоційна втома	Регламентовані перерви кожні 2 години, гімнастика для очей, ергономічні меблі
4	Шум і вібрація	Невеликий рівень шуму від вентиляторів, обладнання	Подразнення, зниження концентрації	Використання тихих систем охолодження, розміщення техніки у звукоізованих корпусах

Особлива увага приділяється мікроклімату приміщення. Оптимальна температура повітря в холодний період року повинна бути в межах 20–22 °С, у теплий – 22–25 °С, відносна вологість – 40–60 %. Рівень шуму не повинен перевищувати 50 дБА, а концентрація пилу та шкідливих речовин у повітрі має відповідати чинним гігієнічним нормам. Для підтримання комфортного мікроклімату використовуються системи кондиціонування та регулярне провітрювання.

У межах діяльності, що пов'язана з обробленням геопросторових даних і супутникових знімків, працівники використовують потужні обчислювальні системи, тому важливо контролювати рівень електромагнітного випромінювання. Всі комп'ютери та мережеве обладнання мають бути заземлені, а електричні розетки – справні й відповідати вимогам електробезпеки. Кабелі прокладаються у спеціальних коробах або каналах, що запобігає їх механічним пошкодженням і випадковим коротким замиканням.

Для запобігання нещасним випадкам проводяться регулярні інструктажі з охорони праці:

- 1) вступний інструктаж – під час прийняття працівника на роботу;
- 2) первинний інструктаж – безпосередньо на робочому місці перед початком виконання завдань;
- 3) повторний інструктаж – не рідше одного разу на шість місяців;
- 4) позаплановий інструктаж – у разі зміни технологічних процесів або оновлення обладнання.

Усі працівники повинні бути ознайомлені з правилами користування первинними засобами пожежогасіння, розташуванням аварійних виходів і діями у разі виникнення надзвичайних ситуацій. У приміщеннях обов'язково мають бути вогнегасники, аптечка першої допомоги, засоби індивідуального захисту (рукавички, окуляри, респіратори – за необхідності).

Серед заходів із покращення умов праці та підвищення рівня безпеки варто виокремити:

- 1) модернізацію робочих місць відповідно до сучасних ергономічних стандартів;
- 2) запровадження системи моніторингу мікроклімату з автоматичним контролем температури, вологості й освітлення;
- 3) використання програмного забезпечення для відстеження навантаження працівників і нагадування про перерви;
- 4) впровадження культури безпечної поведінки та формування відповідального ставлення до охорони праці через навчання й тренінги.

Загалом аналіз умов праці свідчить, що створення безпечного, комфортного та технологічно оснащеного робочого середовища є важливою передумовою ефективного виконання завдань у сфері прогнозування врожайності. Дотримання вимог охорони праці не лише знижує ризики професійних захворювань, а й сприяє підвищенню продуктивності праці, точності аналітичних розрахунків і якості наукових досліджень.

4.2. Розробка логічно-імітаційної моделі травматизму під час монтажу інформаційної системи прогнозування врожайності

Для обґрунтованого аналізу ризиків при монтажі та налагодженні програмно-апаратного комплексу інформаційної системи запропоновано логічно-імітаційну модель, яка поєднує дискретно-подійний підхід із простим марківським описом станів безпеки. Модель призначена показати основні шляхи виникнення нещасних випадків у процесі виконання робіт (монтаж обладнання, прокладання кабелів, робота з переносною технікою, підключення джерел живлення, конфігурація серверів) і оцінити вплив запобіжних заходів на очікувану кількість інцидентів.

У моделі розглядаються три типи агентів: монтажники (фахівці з інсталяції), технічний персонал (адміністратори, електрики) і допоміжний

персонал (кур'єри, лаборанти). Кожний агент у процесі виконання робіт послідовно проходить набір операцій (рис. 4.1).



Рисунок 4.1 – Процес монтажу та налагодження програмно-апаратного комплексу інформаційної системи

Кожна операція має власну тривалість t_i (години) і ймовірність виникнення події «небезпека» p_i , що може перерости в «інцидент» з верогідністю q_i або завершитися «наближеним інцидентом» (near miss) з ймовірністю $1 - q_i$.

Модель формалізується як дискретно-подійний процес з наступними компонентами:

Стани процесу (S)

$$S = \{ Safe, NearMiss, MinorInjury, MajorInjury, Fatality \}. \quad (4.1)$$

Перехід між станами відбувається під впливом подій, що генеруються під час виконання операцій.

Інтенсивність подій

Події «потенційно небезпечна ситуація» моделюються як пуассонівський потік з інтенсивністю λ_i для i -ї операції. Очікувана кількість подій за операцією:

$$E[N_i] = \lambda_i \cdot T_i, \quad (4.2)$$

де T_i – сумарний час виконання операції для всіх агентів.

Ймовірності ескалації

Якщо виникає потенційно небезпечна подія, вона може завершитися у різні стани з умовними ймовірностями.

У практичній реалізації модель імітаційно генерує послідовність операцій для заданої бригади і випадкові події згідно з пуассонівським процесом; далі для кожної події приймається рішення про результат (near miss, minor, major) за вказаними ймовірностями. Модель допускає налаштування параметрів згідно з фактичними умовами: числом робочих бригад, тривалістю монтажних робіт, наявністю освітлення, типом устаткування, наявністю заземлення та ін.

Нижче подано таблицю 4.2 станів та типових параметрів моделі (параметри наведено як демонстраційні – їх необхідно підґрунтовно оцінити на основі місцевих даних і інструктажів).

Таблиця 4.2 – Таблиця станів та параметри імітаційної моделі

Операція	Тип робіт	Середній час на операцію t_i , хв	Інтенсивність подій λ_i , подій/год	Ймовірність Minor, q_i	Ймовірність Major r_i	Примітки
1	Розвантаження обладнання	30	0.05	0.10	0.01	ризик травм при підйомі
2	Переміщення та розміщення	40	0.04	0.08	0.005	можливі підсковзування
3	Прокладання кабелів	60	0.08	0.12	0.02	ризик спотикання, порушення ізоляції
4	Монтаж стійок/шаф	90	0.03	0.06	0.01	падіння інструментів
5	Підключення живлення	20	0.02	0.04	0.02	електробезпека – високий ризик при помилках
6	Налагодження обладнання	120	0.01	0.02	0.005	ментальна втома, помилки оператора

У таблиці наведено значення λ_i та ймовірностей q_i , r_i , які відображають відносну частоту потенційно небезпечних ситуацій та ймовірність їх переростання у травми. Для сценарію без додаткових заходів контролю (базове) імітація на період одного сезону (для 3-особової бригади, сумарний робочий час кожної операції помножений на кількість повторів) дає такі умовні результати:

Таким чином, запропонована логічно-імітаційна модель забезпечує структуровану і контрольовану методику оцінки травматизму під час монтажу інформаційної системи прогнозування врожайності, дозволяє кількісно оцінити ефективність профілактичних заходів та служить інструментом для прийняття управлінських рішень щодо охорони праці.

4.3. Розробка заходів щодо безпеки у надзвичайних ситуаціях

Розроблення та впровадження інформаційної системи прогнозування врожайності потребує врахування можливих ризиків надзвичайних ситуацій, які можуть виникнути як у процесі монтажу, так і під час експлуатації системи. До потенційно небезпечних чинників належать пожежі, відключення електроенергії, затоплення приміщень, ураження електричним струмом, а також інформаційні інциденти, пов'язані з кібератаками або втратою даних. Метою цього підрозділу є формування комплексу заходів, спрямованих на забезпечення безпеки персоналу, збереження інформаційних ресурсів і безперервність роботи системи в умовах надзвичайних подій.

Найпоширенішими джерелами надзвичайних ситуацій у комп'ютеризованих середовищах є порушення правил електробезпеки, коротке замикання, перевантаження електромереж, несправність обладнання, а також людський фактор, який проявляється у недотриманні інструкцій чи помилках під час експлуатації. Значну загрозу становлять природні фактори – грози, буревії, підтоплення – які можуть пошкодити серверне обладнання та

призвести до втрати живлення. Не менш небезпечними є інформаційні ризики, що виникають унаслідок кібератак, збоїв у роботі мережевих сервісів або несанкціонованого доступу до даних. Для запобігання цим загрозам необхідно забезпечити багаторівневу систему захисту, яка поєднує технічні, організаційні та адміністративні заходи.

Особливу увагу слід приділяти захисту інформаційного середовища. Використання антивірусного програмного забезпечення, систем багаторівневої автентифікації, міжмережевих екранів і регулярне оновлення безпекових протоколів знижує ризик кібератак і несанкціонованого доступу. Наявність плану відновлення після аварій (Disaster Recovery Plan) дає змогу оперативно відновити функціонування системи після порушення її роботи – зокрема, визначити пріоритетність запуску серверів, перевірку цілісності даних і тестування працездатності сервісів.

Отже, система безпеки у надзвичайних ситуаціях повинна бути комплексною, передбачати не лише технічний захист, а й підготовку персоналу до швидкого реагування. Реалізація зазначених заходів гарантує зниження ризиків для життя і здоров'я працівників, збереження інформаційних ресурсів та забезпечення стабільного функціонування інформаційної системи прогнозування врожайності навіть у кризових умовах.

РОЗДІЛ 5.

ВИЗНАЧЕННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ОЗИМОЇ ПШЕНИЦІ

Оцінка економічної ефективності впровадження інформаційної системи прогнозування врожайності озимої пшениці є важливим етапом, який дозволяє визначити доцільність інвестицій у розробку, впровадження та експлуатацію програмного комплексу. Головна мета такої оцінки полягає у визначенні, наскільки використання інтелектуальних технологій дає змогу підвищити результативність управління агровиробництвом, зменшити витрати ресурсів та мінімізувати ризики недоотримання урожаю внаслідок погодних коливань.

Для визначення ефективності використання інформаційної системи застосовується показник економічного ефекту E , який обчислюється як різниця між отриманим економічним результатом (додатковим прибутком) та витратами на розробку і впровадження системи:

$$E = (P_1 - P_0) - C_{\text{сист}}, \quad (5.1)$$

де P_1 – прибуток господарства після впровадження системи, грн; P_0 – прибуток до впровадження системи, грн; $C_{\text{сист}}$ – сукупні витрати на створення та обслуговування системи, грн.

Ефективність також доцільно оцінювати за показником рентабельності інвестицій (ROI – Return on Investment):

$$ROI = \frac{E}{C_{\text{сист}}} \times 100\%, \quad (5.2)$$

Якщо значення ROI перевищує 20–25 %, інвестиції вважаються економічно доцільними.

Крім того, можна розрахувати строк окупності проєкту $T_{\text{ок}}$:

$$T_{\text{ок}} = \frac{C_{\text{сист}}}{P_1 - P_0}, \quad (5.3)$$

Чим менше значення $T_{\text{ок}}$, тим швидше система відшкодовує витрати.

У дослідженні розглянуто середнє фермерське господарство площею 500 га, на якому вирощується озима пшениця. Впровадження інформаційної системи прогнозування врожайності дозволило оптимізувати планування агротехнічних операцій (підживлення, зрошення, обробіток посівів), що призвело до підвищення урожайності на 0,45 т/га (з 5,00 до 5,45 т/га). Середня ціна реалізації пшениці становить 7500 грн/т, а питомі виробничі витрати без системи – 28000 грн/га.

Витрати на розробку та впровадження системи (програмне забезпечення, супутникові дані, навчання персоналу, технічне обслуговування) оцінюються у 520 000 грн.

Додатковий обсяг урожаю після впровадження системи:

$$Q_{\text{дод}} = (Y_1 - Y_0) \cdot S = (5.45 - 5.00) \cdot 500 = 225 \text{ т.}$$

Додатковий дохід від реалізації продукції:

$$P_{\text{дод}} = Q_{\text{дод}} \cdot C_{\text{ц}} = 225 \cdot 7500 = 1687500 \text{ грн.}$$

Загальний прибуток після впровадження:

$$P_1 = (Y_1 \cdot S \cdot C_{\text{ц}}) - (S \cdot C_{\text{вир}}) = (5.45 \cdot 500 \cdot 7500) - (500 \cdot 28000) = 9337500 \text{ грн.}$$

До впровадження системи:

$$P_0 = (Y_0 \cdot S \cdot C_{\text{ц}}) - (S \cdot C_{\text{вир}}) = (5.0 \cdot 500 \cdot 7500) - (500 \cdot 28000) = 8500000 \text{ грн.}$$

Економічний ефект:

$$E = (P_1 - P_0) - C_{\text{сист}} = (9,337,500 - 8,500,000) - 520,000 = 317500 \text{ грн.}$$

Рентабельність інвестицій:

$$ROI = \frac{317,500}{520,000} \times 100\% = 61.1\%.$$

Термін окупності:

$$T_{\text{ок}} = \frac{520000}{837500} = 0.62 \text{ року.}$$

Результати розрахунків подано у таблиці 5.1.

Таблиця 5.1 – Результати розрахунку економічної ефективності від використання інформаційної системи прогнозування врожайності озимої пшениці

Показник	Позначення	Значення
Площа посівів	S	500 га
Урожайність до впровадження	Y_0	5 т/га
Урожайність після впровадження	Y_1	5.45 т/га
Ціна реалізації	C_s	7500 грн/т
Витрати на виробництво	C_v	28000 грн/га
Вартість системи	$C_{сист}$	520000 грн
Економічний ефект	E	317500 грн
Рентабельність інвестицій	ROI	61.1 %
Строк окупності	$T_{ок}$	0.62 року

Проведені розрахунки підтверджують, що впровадження інформаційної системи прогнозування врожайності озимої пшениці є економічно доцільним, так як річний економічний ефект становить 317500 грн. Витрати на її розробку та підтримку окупаються менш ніж за один рік, а рівень рентабельності інвестицій перевищує 60 %. Це свідчить про високий потенціал цифрових технологій і глибинних нейронних мереж у підвищенні ефективності сучасного аграрного виробництва.

ВИСНОВКИ І ПРОПОЗИЦІЇ

На підставі проведеного аналізу сучасних наукових підходів встановлено, що ключова тенденція полягає у переході від класичних статистичних методів до глибинних нейронних мереж і гібридних моделей. Це зумовлено потребою врахування складних просторово-часових закономірностей. Пропонується у подальших дослідженнях акцентувати увагу на впровадженні трансформерних архітектур та поєднанні кількох моделей для підвищення точності прогнозів.

Виконано узагальнення біологічних та фенологічних характеристик озимої пшениці. Це дало можливість встановити, що точність прогнозів суттєво залежить від врахування фаз розвитку культури та критичних періодів водозабезпечення. Пропонується при моделюванні використовувати часові ряди вегетаційних індексів, зокрема NDVI та GNDVI, а також інтегрувати агрокліматичні змінні для коректного відображення стану посівів.

На підставі порівняння статистичних, процесно-орієнтованих та інтелектуальних методів доведено, що класичні підходи зручні для інтерпретації, однак поступаються в точності сучасним моделям ШІ. Найвищі результати демонструють CNN, LSTM та їхні модифікації з механізмами уваги. Пропонується розвивати напрям гібридних архітектур та використовувати оптимізаційні алгоритми (IGWO, PSO) для налаштування гіперпараметрів.

Виконаний аналіз практик застосування супутникових даних дав можливість встановити, що найбільш ефективними є комплексні підходи, які поєднують оптичні, радарні й термальні сенсори. Це дозволяє отримати більш повну картину стану агроєкосистем. Пропонується впроваджувати інтегровані платформи моніторингу, які поєднують дані ДЗЗ із геоінформаційними системами та алгоритмами глибинного навчання для оперативного прийняття управлінських рішень у точному землеробстві.

Теоретичні засади прогнозування врожайності сільськогосподарських культур базуються на поєднанні біофізичних моделей, супутникових спостережень та методів глибинного навчання. Використання спектральних

індексів NDVI, EVI, GNDVI та комплексної обробки часових рядів дозволяє суттєво підвищити точність оцінювання продуктивності культур, зокрема озимої пшениці. Поєднання різних джерел даних у рамках нейронних архітектур забезпечує універсальність і адаптивність систем прогнозування для потреб сучасного агровиробництва.

Особливість використання супутникових знімків у агромоніторингу полягає у можливості систематичного спостереження за станом посівів на великих територіях із високою точністю. Поєднання спектральних, радарних і термальних даних забезпечує комплексне розуміння процесів росту культур, а їх інтеграція в нейромережеві моделі дозволяє здійснювати точні прогнози врожайності з урахуванням просторово-часових закономірностей розвитку рослин.

Розвиток глибинних архітектур – від класичних CNN і LSTM до гібридних CNN–LSTM і Transformer – забезпечує суттєвий прогрес у прогнозуванні врожайності. Комбінація супутникових зображень і часових рядів дає змогу створювати адаптивні, високоточні моделі, що враховують як просторові, так і динамічні фактори формування врожаю озимої пшениці.

Вибір інструментальних засобів базується на поєднанні відкритих технологій з високою продуктивністю. Комбінація Python, TensorFlow, Planetary Computer і PostGIS забезпечує повний цикл обробки – від отримання супутникових даних до побудови прогнозу врожайності й його інтерактивної візуалізації. Такий підхід гарантує масштабованість, відтворюваність і практичну придатність розробленої системи в аграрному виробництві.

Нами проведено збір і підготовку вхідних супутникових та метеорологічних даних для Буського району Львівської області. Розроблено скрипт у середовищі Jupyter Notebook для формування 10-денних композитів NDVI та GNDVI, маскуванню хмар і інтеграції з кліматичними параметрами ERA5-Land. Результатом етапу є чистий, узгоджений часовий ряд ознак, придатний для навчання глибинних моделей прогнозування врожайності озимої пшениці.

Проведено багатоспектральний аналіз посівів озимої пшениці на основі супутникових даних Sentinel-2. Створено карти просторового розподілу NDVI

та GNDVI, які демонструють високу неоднорідність розвитку рослинності в межах полів Буського району. Отримані результати підтвердили придатність використання вегетаційних індексів як інформативних предикторів для побудови моделей прогнозування врожайності.

Побудований часовий ряд NDVI та GNDVI для посівів озимої пшениці у межах Буського району Львівської області дозволяє простежити закономірності сезонних змін інтенсивності вегетації. Індекси чітко відображають ключові фази росту – від кущення до дозрівання, що забезпечує їх ефективне застосування як вхідних змінних у задачах прогнозування врожайності за допомогою глибинних нейронних мереж.

Для оцінки ефективності різних архітектур було протестовано п'ять моделей: CNN (1D), LSTM, GRU, CNN–LSTM та Transformer. Кожна з них мала різну складність, кількість параметрів і специфіку роботи з часовими рядами.

Проведена оцінка ефективності моделей підтвердила доцільність застосування архітектури Transformer для прогнозування врожайності озимої пшениці. Її висока точність, гнучкість та здатність інтегрувати різнотипні джерела даних роблять її оптимальним рішенням для сучасних інформаційних систем агроаналітики, що базуються на супутникових спостереженнях і метеоданих.

Розроблено заходи із охорони праці. Створення безпечного, комфортного та технологічно оснащеного робочого середовища є важливою передумовою ефективного виконання завдань у сфері прогнозування врожайності. Дотримання вимог охорони праці не лише знижує ризики професійних захворювань, а й сприяє підвищенню продуктивності праці.

Проведені розрахунки підтверджують, що впровадження інформаційної системи прогнозування врожайності озимої пшениці є економічно доцільним, так як річний економічний ефект становить 317500 грн. Витрати на її розробку та підтримку окупаються менш ніж за один рік, а рівень рентабельності інвестицій перевищує 60%. Це свідчить про високий потенціал цифрових технологій і глибинних нейронних мереж у підвищенні ефективності сучасного аграрного виробництва.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Жидецький В.Ц., Джигирей В.С., Мельников О.В. Основи охорони праці. Підручник. Вид. 5-е, доповнене. Львів: Афіша, 2012. 350с.
2. Класифікація в Python з Scikit-Learn та Pandas. URL: <https://stackabuse.com/classification-in-python-with-scikit-learn-and-pandas/> (дата звернення: 14.09.2025).
3. Лехман С.Д., Рублев В.І., Рябцев Б.І. Запобігання аварійності і травматизму у сільському господарстві. К.: Урожай, 1993. 267 с.
4. Ситнік Б.Т. Основи інформаційних систем і технологій: навч. посіб. Харків: УкрДУЗТ, 2018. 130 с.
5. Тригуба А., Маланчук О., Тригуба І., Мармуляк А., Демчина В. Андрушків О., Олійник Р. Вплив сучасних інформаційних технологій на процеси ініціації та планування проєктів розвитку громад та регіонів. Вісник Львівського національного університету природокористування: агроінженерні дослідження. №24. Львів: Львів НУП, 2024. С. 123-130.
6. Грицюк П., Бабич Т., Красько Б. Класифікаційні методи прогнозування врожайності. Вісник Хмельницького національного університету. Технічні науки. 2022. Т. 309, № 3. С. 209–216. DOI: <https://doi.org/10.31891/2307-5732-2022-309-3-209-216>
7. Muruganantham P., Wibowo S., Grandhi S., Samrat N. H., Islam N. A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing. Remote Sens. 2022. Т. 14, № 9. Стр. 1990. DOI: <https://doi.org/10.3390/rs14091990>
8. Ashfaq M., Khan I., Shah D., Ali S., Tahir M. Predicting wheat yield using deep learning and multi-source environmental data. Scientific Reports. 2025. Vol. 15. Article number: 26446. DOI: <https://doi.org/10.1038/s41598-025-26446-7>
9. Kogan F., Kussul N., Adamenko T., Skakun S., Kravchenko O., Kryvobok O., Shelestov A., Kolotii A., Kussul O., Lavrenyuk A. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and

biophysical models. *International Journal of Applied Earth Observation and Geoinformation*. 2013. Vol. 23. P. 192–203. DOI: <https://doi.org/10.1016/j.jag.2013.01.002>

10. Battude M., Al Bitar A., Morin D., Cros J., Huc M., Sicre C., Le Dantec V. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 data. *Remote Sensing of Environment*. 2016. Vol. 184. P. 668–681. <https://doi.org/10.1016/j.rse.2016.07.030>

11. Khaki S., Wang L. Crop yield prediction using deep neural networks. *Frontiers in Plant Science*. 2019. Vol. 10. P. 621. DOI: <https://doi.org/10.3389/fpls.2019.00621>

12. Sun J., Di L., Sun Z., Shen Y., Lai Z., Rahman M. Hybrid CNN–RNN for remote sensing image classification. *Remote Sensing*. 2020. Vol. 12, No. 19. P. 3085. DOI: <https://doi.org/10.3390/rs12193085>

13. Kogan F., Kussul N., Adamenko T., Skakun S., Kravchenko O., Kryvobok O., Shelestov A., Kolotii A., Kussul O., Lavrenyuk A. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. *International Journal of Applied Earth Observation and Geoinformation*. 2013. Vol. 23. P. 192–203. DOI: <https://doi.org/10.1016/j.jag.2013.01.002>

14. Fan J., Zhang F., Yang R., Chen P., Xia X., Jiang Y. Attention-based deep learning model for wheat yield prediction using multi-source data. *Computers and Electronics in Agriculture*. 2022. Vol. 198. P. 107091. DOI: <https://doi.org/10.1016/j.compag.2022.107091>

15. Kuwata K., Shibasaki R. Estimating crop yields with deep learning and remotely sensed data. *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2015. P. 858–861. DOI: <https://doi.org/10.1109/IGARSS.2015.7325888>

16. Lundberg S. M., Lee S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems (NIPS 2017)*. 2017. Vol. 30. P. 4765–4774.

17. Nebiker S., Lack N., Abächerli M., Läderach S. Light-weight multispectral UAV sensors and their capabilities for predicting grain yield and detecting plant diseases. *International Journal of Remote Sensing*. 2020. Vol. 41, No. 19. P. 7494–7515. DOI: <https://doi.org/10.1080/01431161.2020.1752950>
18. Joshi A., Pradhan B., Gite S., Chakraborty S. Remote-Sensing Data and Deep-Learning Techniques in Crop Mapping and Yield Prediction: A Systematic Review. *Remote Sensing*. 2023. Vol. 15, No. 8. P. 2014. DOI: <https://doi.org/10.3390/rs15082014>
19. Jędrejek A., Pudełko R. Exploring the potential use of Sentinel-1 and 2 satellite imagery for monitoring winter wheat growth under agricultural drought conditions in North-Western Poland. *Agriculture*. 2023. Vol. 13, No. 9. P. 1798. DOI: <https://doi.org/10.3390/agriculture13091798>
20. Lobell D. B., Asseng S. Comparing estimates of climate change impacts from process-based and statistical crop models. *Environmental Research Letters*. 2017. Vol. 12, No. 1. P. 015001. DOI: <https://doi.org/10.1088/1748-9326/aa518a>
21. Pu Q., Xi Z., Yin S., Zhao Z. Advantages of transformer and its application for medical image segmentation: a survey. *BioMedical Engineering OnLine*. 2024. Vol. 23, No. 1. P. 12. DOI: <https://doi.org/10.1186/s12938-024-01212-4>
22. Justice C., Becker-Reshef I., Vermote E., Privette J., Masuoka E., Wolfe R., Claverie M., Nagol J. Developing a strategy for global agricultural monitoring in the framework of GEO. *Photogrammetric Engineering & Remote Sensing*. 2011. Vol. 77, No. 12. P. 123–137. DOI: <https://doi.org/10.14358/PERS.77.12.123>
23. Veloso A., Mermoz S., Bouvet A., Le Toan T., Planells M., Dejoux J. F., Ceschia E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*. 2017. Vol. 199. P. 415–426. DOI: <https://doi.org/10.1016/j.rse.2017.07.015>
24. Kogan F. N. Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*. 1995. Vol. 15, No. 11. P. 91–100. DOI: [https://doi.org/10.1016/0273-1177\(95\)00079-T](https://doi.org/10.1016/0273-1177(95)00079-T)

25. Skakun S., Vermote E., Roger J. C., Franch B., Kussul N., Ju J., Kogan F., Justice C. Early season large-area winter crop mapping using MODIS NDVI data, growing degree days information and a Gaussian mixture model. *Remote Sensing of Environment*. 2017. Vol. 195. P. 36–46. DOI: <https://doi.org/10.1016/j.rse.2017.04.026>
26. Jędrejek A., Pudełko R. Exploring the potential use of Sentinel-1 and 2 satellite imagery for monitoring winter wheat growth under agricultural drought conditions in North-Western Poland. *Agriculture*. 2023. Vol. 13, No. 9. P. 1798. DOI: <https://doi.org/10.3390/agriculture13091798>
27. Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, 199, 415–426. DOI: <https://doi.org/10.1016/j.rse.2017.07.015>
28. Battude M. et al. Estimating maize biomass and yield over large areas using Sentinel-2 data. *Remote Sensing of Environment*. 2016. Vol. 184. P. 668–681. DOI: <https://doi.org/10.1016/j.rse.2016.07.030>
29. Hochreiter S., Schmidhuber J. Long Short-Term Memory. *Neural Computation*. 1997. 9(8):1735–1780.
30. Choi, S. R., & Lee, M. (2023). Transformer Architecture and Attention Mechanisms in Genome Data Analysis: A Comprehensive Review. *Biology*, 12(7), 1033. DOI: <https://doi.org/10.3390/biology12071033>
31. Chollet F. *Deep Learning with Python*. Manning Publications, 2021.
32. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint arXiv:1412.3555*. DOI: <https://doi.org/10.48550/arXiv.1412.3555>
33. Kattenborn, T., Eichel, J., & Fassnacht, F. E. (2019). Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Scientific Reports*, 9, Article 17656. DOI: <https://doi.org/10.1038/s41598-019-53797-9>

ДОДАТКИ

Додаток А.1

Код підготовки вхідних даних дистанційного зондування і метеопараметрів

```
# ===== 0) Спільні параметри: AOI та період 2024 =====
import json, pandas as pd, numpy as np
from shapely.geometry import shape
aoi_bbox = box(24.55, 49.88, 24.82, 50.05) # з ТВОГО п.1
aoi_geojson = {
    "type": "FeatureCollection", "features": [{"type": "Feature", "geometry": mapping(aoi_bbox), "properties": {}}]}
start_date = "2024-03-01"
end_date = "2024-07-31"

# 10-денні вікна
TEN_DAYS = pd.date_range(start=start_date, end=end_date, freq="10D")
TEN_DAYS = list(zip(TEN_DAYS[:-1], TEN_DAYS[1:]))

# ===== 1) S2: пошук і завантаження з обрізанням по AOI =====
!pip -q install planetary-computer pystac-client odc-stac rioxarray xarray matplotlib folium

import planetary_computer as pc
from pystac_client import Client
from odc.stac import load as stac_load
import xarray as xr, numpy as np, matplotlib.pyplot as plt

stac = Client.open("https://planetarycomputer.microsoft.com/api/stac/v1")
search = stac.search(
    collections=["sentinel-2-l2a"],
    intersects=aoi_geojson["features"][0]["geometry"],
    datetime=f"{start_date}/{end_date}",
    query={"eo:cloud_cover":{"lt":60}}
)
items = list(search.items())
items = [pc.sign(i) for i in items]
print("S2 scenes:", len(items))

data = stac_load(
    items,
    bands=["B02", "B03", "B04", "B08", "SCL"],
    geopolygon=shape(aoi_geojson["features"][0]["geometry"]), # <-- точне AOI
    crs="EPSG:3857",
    resolution=60, # швидко і стабільно
    chunks={"time":1, "x":1024, "y":1024},
    fail_on_error=False
)

# маска хмар SCL
cloud_classes = [3,8,9,10,11]
cloud_mask = ~data["SCL"].isin(cloud_classes)

B03 = xr.where(cloud_mask, data["B03"], np.nan).astype("float32")
```

```

B04 = xr.where(cloud_mask, data["B04"], np.nan).astype("float32")
B08 = xr.where(cloud_mask, data["B08"], np.nan).astype("float32")

ndvi = (B08 - B04) / (B08 + B04)
gndvi = (B08 - B03) / (B08 + B03)
ndvi.name, gndvi.name = "NDVI", "GNDVI"

# ===== 2) 10-денні композити: медіана по часу -> середнє по площі =====
def spatial_mean(da): return da.mean(dim=("x", "y"), skipna=True)

rows = []
time_index = ndvi["time"].values
for t0, t1 in TEN_DAYS:
    sel = (time_index >= np.datetime64(t0)) & (time_index < np.datetime64(t1))
    if sel.sum() == 0:
        continue
    ndvi_med = ndvi.sel(time=time_index[sel]).median(dim="time", skipna=True)
    gndvi_med = gndvi.sel(time=time_index[sel]).median(dim="time", skipna=True)
    rows.append({
        "t0": pd.to_datetime(t0),
        "t1": pd.to_datetime(t1),
        "NDVI_mean": float(spatial_mean(ndvi_med).values),
        "GNDVI_mean": float(spatial_mean(gndvi_med).values),
        "obs_count": int(sel.sum())
    })

ts_veg = pd.DataFrame(rows).sort_values("t0").reset_index(drop=True)
ts_veg.to_csv("veg_10d.csv", index=False)
print("veg_10d.csv saved"); display(ts_veg.head())

# ===== 3) ERA5-Land через Open-Meteo (2024), місячні -> 10-денки =====
import requests, time as tm
from shapely.geometry import Point, Polygon
coords = [(24.55, 49.88), (24.82, 49.88), (24.82, 50.05), (24.55, 50.05), (24.55, 49.88)]
poly = Polygon(coords)

lons = np.arange(24.55, 24.82+1e-9, 0.05)
lats = np.arange(49.88, 50.05+1e-9, 0.05)
pts = [(lo, la) for lo in lons for la in lats if poly.buffer(1e-8).contains(Point(lo, la))]
if len(pts) > 25:
    idx = np.round(np.linspace(0, len(pts)-1, 25)).astype(int)
    pts = [pts[i] for i in idx]

def om_monthly(lon, lat):
    url = "https://archive-api.open-meteo.com/v1/era5"
    params = {
        "latitude": lat, "longitude": lon,
        "start_date": start_date, "end_date": end_date,
        "models": "era5_land",
        "monthly": "temperature_2m_mean,precipitation_sum,et0_fao_evapotranspiration_sum",
        "timezone": "UTC"
    }

```

```

r = requests.get(url, params=params, timeout=60); r.raise_for_status()
m = r.json().get("monthly", {})
t = m.get("time", [])
if not t: return None
df = pd.DataFrame({"date":pd.to_datetime(t)})
if "temperature_2m_mean" in m:
df["temperature_C"] = np.array(m["temperature_2m_mean"], float)
if "precipitation_sum" in m: df["precip_mm"] = np.array(m["precipitation_sum"], float)
if "et0_fao_evapotranspiration_sum" in m:
df["et0_mm"] = np.array(m["et0_fao_evapotranspiration_sum"], float)
return df

dfs=[]
for lo,la in pts:
    try:
        dfp = om_monthly(lo,la)
        if dfp is not None and not dfp.empty:
            dfs.append(dfp)
    except Exception as e:
        continue

if len(dfs)==0:
    # fallback: центрoїд
    c = poly.centroid; dfp = om_monthly(float(c.x), float(c.y))
    dfs = [dfp]

era = (pd.concat(dfs, ignore_index=True)
        .groupby("date")[["temperature_C", "precip_mm", "et0_mm"]]
        .mean().reset_index().sort_values("date"))
era.to_csv("era_monthly.csv", index=False)
print("era_monthly.csv saved"); display(era.head())

# місячні -> 10-денки (інтерполяція по часу)
era_10d = (era.set_index("date").resample("10D").interpolate("time").reset_index())
era_10d.to_csv("era_10d.csv", index=False)
print("era_10d.csv saved"); display(era_10d.head())

# ===== 4) Злиття S2 (10D) + ERA (10D) і контроль якості =====
dataset = pd.merge_asof(
    ts_veg.rename(columns={"t0":"date"}).sort_values("date"),
    era_10d.sort_values("date"),
    on="date"
)
# базові перевірки
missing = dataset.isna().mean().sort_values(ascending=False)
print("Пропуски, частка:"); print(missing.head(10))

dataset.to_csv("dataset_10d.csv", index=False)
print("dataset_10d.csv saved"); display(dataset.head())

```

Додаток А.2

Код системи прогнозування врожайності

```
# tk_app.py
# Tkinter-інтерфейс без PySide6. Роки 2019–2024.

import os, json, tkinter as tk
from tkinter import ttk, filedialog, messagebox
from pathlib import Path
import numpy as np
import pandas as pd
import matplotlib
matplotlib.use("TkAgg")
import matplotlib.pyplot as plt
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg

TF_OK = False
try:
    from tensorflow import keras
    TF_OK = True
except Exception:
    TF_OK = False

DEFAULT_YEARS = (2019, 2020, 2021, 2022, 2023, 2024)

# ----- utils -----
def read_features_folder(data_dir: str, years=DEFAULT_YEARS) -> pd.DataFrame:
    paths = []
    for y in years:
        p = Path(data_dir) / f"features_busky_{y}_10d.csv"
        if p.exists(): paths.append(str(p))
    if not paths: return pd.DataFrame()
    dfs = []
    for p in paths:
        try:
            df = pd.read_csv(p)
            if "date" in df.columns: df["date"] = pd.to_datetime(df["date"])
            elif "t0" in df.columns: df["date"] = pd.to_datetime(df["t0"])
            else:
                df["date"] = pd.to_datetime("2024-01-01") + pd.to_timedelta(np.arange(len(df)),
unit="D")
            for col in ["NDVI_mean", "GNDVI_mean", "temperature_C", "precip_mm", "et0_mm", "year", "month"]:
                if col not in df.columns: df[col] = np.nan
            if df["year"].isna().all():
                try:
                    y_infer = int(Path(p).stem.split("_")[2])
                    df["year"] = y_infer
                except Exception:
                    df["year"] = df["date"].dt.year
```

```

        dfs.append(df)
    except Exception as e:
        print("Помилка читання", p, e)
    if not dfs: return pd.DataFrame()
    all_df = pd.concat(dfs, ignore_index=True).sort_values(["year", "date"]).reset_index(drop=True)
    for c in ["temperature_C", "precip_mm", "et0_mm"]:
        if c in all_df.columns:
            all_df[c]=all_df.groupby("year")[c].apply(lambda
s.interpolate(limit_direction="both").ffill().bfill()).values
            if np.all(pd.isna(all_df[c].values)):
                all_df[c]=15.0 if c=="temperature_C" else 0.0
        if "NDVI_mean" in all_df.columns:
            all_df["NDVI_mean"]=all_df.groupby("year")["NDVI_mean"].apply(
lambda
s:
s.interpolate(limit_direction="both").ffill().bfill().rolling(2,
min_periods=1).mean()
).values
    return all_df

def ndvi_auc(df_year: pd.DataFrame)->float:
    if df_year.empty or "date" not in df_year.columns or "NDVI_mean" not in df_year.columns:
    return np.nan
    d = df_year.dropna(subset=["date", "NDVI_mean"]).sort_values("date")
    if len(d)<2: return np.nan
    t=(d["date"]-d["date"].min()).dt.total_seconds()/(24*3600.0)
    y=d["NDVI_mean"].clip(0,1.0).values
    return float(np.trapz(y,t))

def default_auc_to_yield(auc: float, alpha=0.055, beta=3.2)->float:
    return float(alpha*auc+beta) if not pd.isna(auc) else np.nan

def make_feature_vector_for_model(df_year: pd.DataFrame)->np.ndarray:
    auc=ndvi_auc(df_year)
    t_mean=float(np.nanmean(df_year["temperature_C"])) if "temperature_C" in df_year else np.nan
    p_sum=float(np.nansum(df_year["precip_mm"])) if "precip_mm" in df_year else np.nan
    e_sum=float(np.nansum(df_year["et0_mm"])) if "et0_mm" in df_year else np.nan
    x=np.array([auc,t_mean,p_sum,e_sum], dtype="float32")
    x=np.nan_to_num(x, nan=0.0, posinf=0.0, neginf=0.0)
    return x.reshape(1,-1)

def find_model_file(path_like: str)->Path|None:
    p=Path(path_like)
    if p.is_file() and p.suffix.lower() in [".keras", ".h5"]: return p
    if p.is_dir():
        k=sorted(p.glob("*.keras"))
        if k: return k[0]
        h=sorted(p.glob("*.h5"))
        if h: return h[0]
    return None

# ----- Tk App -----
class App(tk.Tk):
    def __init__(self):

```

```

super().__init__()
self.title("NDVI → Yield (2019–2024) – Tk")
self.geometry("1200x780")

self.data_dir = tk.StringVar(value=r"F:\Dipl_2025\Mar_2025\Тирпак\Data")
self.model_path = tk.StringVar(value="")
self.alpha = tk.DoubleVar(value=0.055)
self.beta = tk.DoubleVar(value=3.2)

self.df_all = pd.DataFrame()
self.df_year = pd.DataFrame()
self.years_avail = []
self.loaded_model = None
self.used_model_file = None

# верхня панель
top = ttk.Frame(self); top.pack(fill="x", padx=8, pady=6)
ttk.Label(top, text="Тека з даними:").pack(side="left")
ttk.Entry(top, textvariable=self.data_dir, width=60).pack(side="left", padx=4)
ttk.Button(top, text="Обрати...", command=self.choose_dir).pack(side="left", padx=4)
ttk.Button(top, text="Завантажити", command=self.load_data).pack(side="left", padx=4)

ttk.Label(top, text=" Модель:").pack(side="left", padx=(12,2))
ttk.Entry(top, textvariable=self.model_path, width=50).pack(side="left", padx=4)
ttk.Button(top, text="Обрати...", command=self.choose_model).pack(side="left", padx=4)
ttk.Button(top, text="Завантажити модель", command=self.load_model).pack(side="left",
padx=4)

# параметри формули
pfrm = ttk.Frame(self); pfrm.pack(fill="x", padx=8, pady=4)
ttk.Label(pfrm, text="alpha:").pack(side="left")
ttk.Entry(pfrm, textvariable=self.alpha, width=8).pack(side="left", padx=4)
ttk.Label(pfrm, text="beta:").pack(side="left")
ttk.Entry(pfrm, textvariable=self.beta, width=8).pack(side="left", padx=4)

# вибір року + кнопки
mid = ttk.Frame(self); mid.pack(fill="x", padx=8, pady=4)
ttk.Label(mid, text="Рік:").pack(side="left")
self.year_cb = ttk.Combobox(mid, state="readonly", width=10)
self.year_cb.pack(side="left", padx=4)
self.year_cb.bind("<<ComboboxSelected>>", lambda e: self.change_year())
ttk.Button(mid, text="Прогноз", command=self.predict).pack(side="left", padx=6)
ttk.Button(mid, text="Експорт CSV", command=self.export_csv).pack(side="left", padx=6)

# текстові поля
self.txt = tk.Text(self, height=6)
self.txt.pack(fill="x", padx=8, pady=4)

# графіки
figs = ttk.Frame(self); figs.pack(fill="both", expand=True, padx=8, pady=6)
self.fig1, self.ax1 = plt.subplots(figsize=(5.5,3.0))

```

```

        self.can1 = FigureCanvasTkAgg(self.fig1, master=figs);
self.can1.get_tk_widget().pack(side="left", fill="both", expand=True)
        self.fig2, self.ax2 = plt.subplots(figsize=(5.5,3.0))
        self.can2 = FigureCanvasTkAgg(self.fig2, master=figs);
self.can2.get_tk_widget().pack(side="left", fill="both", expand=True)
        self.fig3, self.ax3 = plt.subplots(figsize=(5.5,3.0))
        self.can3 = FigureCanvasTkAgg(self.fig3, master=figs);
self.can3.get_tk_widget().pack(side="left", fill="both", expand=True)

def choose_dir(self):
    d = filedialog.askdirectory(initialdir=self.data_dir.get())
    if d: self.data_dir.set(d)

def choose_model(self):
    fp = filedialog.askopenfilename(filetypes=[("Keras model", ".keras .h5"), ("All files", "*.*)"])
    if fp:
        self.model_path.set(fp)
    else:
        d = filedialog.askdirectory()
        if d: self.model_path.set(d)

def load_model(self):
    if not TF_OK:
        messagebox.showwarning("TensorFlow", "TensorFlow/Keras недоступний.")
        return
    mp = self.model_path.get().strip()
    if not mp:
        messagebox.showinfo("Шлях до моделі", "Вкажіть файл *.keras/*.h5 або теку."); return
    mf = find_model_file(mp)
    if not mf:
        messagebox.showwarning("Моделі", "Не знайдено *.keras/*.h5 за вказаним шляхом.");
return
    try:
        self.loaded_model = keras.models.load_model(mf)
        self.used_model_file = str(mf)
        messagebox.showinfo("Моделі", f"Завантажено: {mf.name}")
    except Exception as e:
        self.loaded_model=None; self.used_model_file=None
        messagebox.showerror("Помилка моделі", str(e))

def load_data(self):
    d = self.data_dir.get().strip()
    df = read_features_folder(d, DEFAULT_YEARS)
    if df.empty:
        messagebox.showerror("Дані", "Не знайдено файлів features_busky_YYYY_10d.csv.")
        return
    self.df_all = df
    self.years_avail = sorted(self.df_all["year"].dropna().unique().astype(int).tolist())
    self.year_cb["values"] = self.years_avail
    self.year_cb.set(self.years_avail[-1])
    self.change_year()
    messagebox.showinfo("Дані", "Завантажено")

```

```

def change_year(self):
    if not self.years_avail: return
    y = int(self.year_cb.get())
    self.df_year = self.df_all[self.df_all["year"]==y].copy()
    self.draw_plots()

def draw_plots(self):
    if self.df_year.empty: return
    # 1) NDVI сезон
    self.ax1.clear()
    self.ax1.plot(self.df_year["date"], self.df_year["NDVI_mean"], marker="o", linewidth=2)
    self.ax1.set_xlabel("Дата"); self.ax1.set_ylabel("NDVI"); self.ax1.set_title("NDVI протягом сезону"); self.ax1.grid(True, alpha=0.25)
    self.can1.draw()
    # 2) Порівняння
    self.ax2.clear()
    for y in self.years_avail:
        sub = self.df_all[self.df_all["year"]==y]
        self.ax2.plot(sub["date"], sub["NDVI_mean"], marker=".", linewidth=1.6, label=str(y))
    self.ax2.legend(ncols=3, fontsize=8); self.ax2.set_xlabel("Дата");
self.ax2.set_ylabel("NDVI"); self.ax2.set_title("Порівняння NDVI за роками");
self.ax2.grid(True, alpha=0.25)
    self.can2.draw()
    # 3) Теплова карта
    self.ax3.clear()
    d = self.df_all.copy(); d["month"]=d["date"].dt.month
    pt = d.pivot_table(index="year", columns="month", values="NDVI_mean", aggfunc="mean")
    im = self.ax3.imshow(pt.values, aspect="auto", origin="upper")
    self.ax3.set_yticks(range(len(pt.index))); self.ax3.set_yticklabels(pt.index.astype(int).tolist())
    self.ax3.set_xticks(range(len(pt.columns)));
self.ax3.set_xticklabels(pt.columns.astype(int).tolist())
    self.ax3.set_xlabel("Місяць"); self.ax3.set_ylabel("Рік"); self.ax3.set_title("Сезонна мапа NDVI")
    for i in range(pt.shape[0]):
        for j in range(pt.shape[1]):
            v = pt.values[i,j]
            if not np.isnan(v):
                self.ax3.text(j,i,f"{v:.2f}", ha="center", va="center", fontsize=8, color="white")
    self.can3.draw()

def predict(self):
    if self.df_year.empty:
        messagebox.showwarning("Прогноз", "Немає даних."); return
    auc = ndvi_auc(self.df_year)
    y_source = "формула"
    y_pred = default_auc_to_yield(auc, self.alpha.get(), self.beta.get())
    if self.loaded_model is not None:
        try:
            X = make_feature_vector_for_model(self.df_year)
            y_pred = float(self.loaded_model.predict(X, verbose=0).ravel()[0])
            y_source = f"модель ({Path(self.used_model_file).name})"

```



```

except Exception as e:
    y_source = f"формула (fallback, помилка моделі: {e})"
    ci = 0.25 + 0.05*max(y_pred,0)
    txt = f"AUC(NDVI): {auc:.2f}\nПрогноз: {y_pred:.2f} т/га\nІнтервал  $\approx$  [{y_pred-ci:.2f}; {y_pred+ci:.2f}] т/га\nДжерело: {y_source}\nФормула: Yield = {self.alpha.get()} * AUC + {self.beta.get()}"
    self.txt.delete("1.0", "end"); self.txt.insert("end", txt)

def export_csv(self):
    if self.df_year.empty:
        messagebox.showwarning("Експорт", "Спершу завантажте дані."); return
    y = int(self.year_cb.get())
    auc = ndvi_auc(self.df_year)
    y_pred = default_auc_to_yield(auc, self.alpha.get(), self.beta.get())
    if self.loaded_model is not None:
        try:
            X = make_feature_vector_for_model(self.df_year)
            y_pred = float(self.loaded_model.predict(X, verbose=0).ravel()[0])
            src = f"модель ({Path(self.used_model_file).name})"
        except Exception:
            src = "формула (fallback)"
    else:
        src = "формула"
    out = pd.DataFrame({"year": [y], "auc": [auc], "yield_pred_t_ha": [y_pred], "source": [src]})
    fp = filedialog.asksaveasfilename(defaultextension=".csv", initialfile=f"summary_{y}.csv")
    if fp:
        out.to_csv(fp, index=False)
        messagebox.showinfo("Експорт", f"Збережено: {fp}")

if __name__ == "__main__":
    App().mainloop()

```