

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ПРИРОДОКОРИСТУВАННЯ**

**ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ  
ТЕХНОЛОГІЙ  
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

# **КВАЛІФІКАЦІЙНА РОБОТА**

першого (бакалаврського) рівня вищої освіти

на тему: «Інформаційна технологія прогнозування урожайності  
сільськогосподарських культур»

Виконав: студент 4 курсу групи Іт-41сп

Спеціальності 126 «Інформаційні системи та  
технології»

(шифр і назва)

Сліпець Дмитро Тарасович

(Прізвище та ініціали)

Керівник: д.т.н., професор Тригуба А.М.

(Прізвище та ініціали)

Рецензент: к.т.н., доцент Бабич М.І.

(Прізвище та ініціали)

**ДУБЛЯНИ-2024**

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

Перший (бакалаврський) рівень вищої освіти  
Спеціальність 126 «Інформаційні системи та технології»

«ЗАТВЕРДЖУЮ»

Завідувач кафедри \_\_\_\_\_

д.т.н., проф. А. М. Тригуба

«\_\_\_\_» \_\_\_\_\_ 2024 р.

## ЗАВДАННЯ

на кваліфікаційну роботу студенту

Сліпцю Дмитру Тарасовичу

1. Тема роботи: «Інформаційна технологія прогнозування урожайності сільськогосподарських культур»

Керівник роботи Тригуба Анатолій Миколайович, професор  
затверджені наказом по університету від 27.11.2023 року № 641/к-с.

2. Строк подання студентом роботи 15.06.2024 р.

3. Вихідні дані до роботи: вимоги до процесу прогнозування урожайності сільськогосподарських культур; засоби розробки, вимоги до технічного забезпечення, архітектура програмного забезпечення, методика прогнозування урожайності сільськогосподарських культур.

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

Вступ.

1. Аналіз предметної області.

2. Проектування інформаційної технології прогнозування урожайності сільськогосподарських культур.

3. Збір та попередня обробка даних для навчання моделі прогнозування урожайності сільськогосподарських культур.

4. Створення та використання моделі прогнозування урожайності сільськогосподарських культур.

5. Охорона праці.

Висновки та пропозиції.

Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових креслень): особливості прогнозування урожайності сільськогосподарських культур; сучасні інформаційні технології для прогнозування урожайності сільськогосподарських культур; проектування інформаційної технології прогнозування урожайності сільськогосподарських культур; збір та попередня обробка даних для навчання моделі прогнозування урожайності сільськогосподарських культур; створення та використання моделі прогнозування урожайності сільськогосподарських культур.

## 6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1, 2, 3, 4	Тригуба А.М., зав. кафедри ІТ		
5	Городецький І.М., доцент кафедри фізики, інженерної механіки та безпеки виробництва		

7. Дата видачі завдання

27 листопада 2023 р.

## Календарний план

№ з/п	Назва етапів кваліфікаційної роботи	Терміни виконання етапів роботи	Примітка
1	Написання першого розділу	27.11-31.12.23	
2	Виконання другого розділу та аркушів ілюстраційного матеріалу до нього	01.02-05.03.24	
3.	Виконання третього розділу та аркушів ілюстраційного матеріалу до нього	06.03-14.04.24	
4.	Виконання четвертого розділу та аркушів ілюстраційного матеріалу до нього	15.04-10.05.24	
5.	Написання розділу «Охорона праці»	11.05-23.05.24	
6.	Завершення оформлення розрахунково-пояснювальної записки та аркушів ілюстраційного матеріалу	24-31.05.24	
7.	Завершення роботи в цілому	01-10.06.24	

Студент \_\_\_\_\_ Сліпець Д.Т.  
(підпис)Керівник роботи \_\_\_\_\_ Тригуба А.М.  
(підпис)

УДК 004.9:631.5

Інформаційна технологія прогнозування урожайності сільськогосподарських культур.

Сліпець Д.Т. Кафедра ІТ – Дубляни, Львівський НУП, 2024.

Кваліфікаційна робота: 80с. текст. част., 18 рис., 5 табл., 12 арк. ілюстраційного матеріалу, 32 джерела.

Подано особливості прогнозування урожайності сільськогосподарських культур. Проаналізовано особливості оцінки врожайності сільськогосподарських культур за допомогою різних підходів. Наведено сучасні інформаційні технології для прогнозування урожайності сільськогосподарських культур. Сформульовано ідею проекту.

Представлена архітектура інформаційної технології прогнозування урожайності сільськогосподарських культур. Здійснено аналіз вимог до інформаційної технології. Вибрано засоби та інструментів. Здійснено проектування інтерфейсу веб-додатку користувачів. Подано інтеграційний інтерфейс (API).

Проаналізовано набір даних для прогнозування урожайності сільськогосподарських культур. Здійснено підготовку даних для навчання моделі прогнозування урожайності сільськогосподарських культур. Подано результати аналізу впливу кліматичних умов на урожайність сільськогосподарських культур.

Здійснено кодування категоріальних змінних. Проведено навчання моделей для прогнозування урожайності сільськогосподарських культур. Оцінено процес навчання моделі RandomForestRegressor та точності прогнозування урожайності сільськогосподарських культур.

Розроблено заходи та інструкцію з охорони праці під час створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур.

## ЗМІСТ

ВСТУП .....	7
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	9
1.1. Особливості прогнозування урожайності сільськогосподарських культур	9
1.2. Оцінка врожайності сільськогосподарських культур за допомогою різних підходів.....	12
1.3. Використання машинного навчання для прогнозування урожайності сільськогосподарських культур .....	18
1.4. Сучасні інформаційні технології для прогнозування урожайності сільськогосподарських культур .....	22
1.5. Ідентифікація ідеї проекту.....	27
РОЗДІЛ 2. ПРОЕКТУВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР .....	29
2.1. Архітектура інформаційної технології прогнозування урожайності сільськогосподарських культур .....	29
2.2. Аналіз вимог до інформаційної технології.....	36
2.3. Вибір засобів та інструментів .....	38
2.4. Проектування інтерфейсу веб-додатку користувачів.....	41
2.5. Інтеграційний інтерфейс (API) .....	43
РОЗДІЛ 3. ЗБІР ТА ПОПЕРЕДНЯ ОБРОБКА ДАНИХ ДЛЯ НАВЧАННЯ МОДЕЛІ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР .....	46
3.1. Набір даних для прогнозування урожайності сільськогосподарських культур.....	46
3.2. Підготовка даних для навчання моделі прогнозування урожайності сільськогосподарських культур .....	47
3.3. Результати аналізу впливу кліматичних умов на урожайність сільськогосподарських культур .....	48

РОЗДІЛ 4. СТВОРЕННЯ ТА ВИКОРИСТАННЯ МОДЕЛІ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР .....	52
4.1. Кодування категоріальних змінних.....	52
4.2. Навчання моделей для прогнозування урожайності сільськогосподарських культур.....	53
4.3. Оцінення процесу навчання моделі RandomForestRegressor та точності прогнозування урожайності сільськогосподарських культур .....	57
РОЗДІЛ 5. ОХОРОНА ПРАЦІ .....	59
5.1. Вимоги до мікроклімату у приміщенні.....	59
5.2. Вимоги до освітлення приміщень та робочих місць .....	60
5.3. Аналіз стану шуму у приміщенні .....	62
5.4. Інструкція з охорони праці під час створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур.....	64
ВИСНОВКИ І ПРОПОЗИЦІЇ.....	67
ДОДАТКИ.....	76
Фрагмент синтаксису інтерфейсу користувача.....	77
Головна панель (HTML) .....	77
Форма введення даних (HTML).....	78
Результати прогнозування (HTML).....	78
Аналітика (HTML) .....	78
Звіти (HTML)мил .....	79
Код реалізації API на Flask (Python).....	79

## ВСТУП

Сільське господарство є однією з найважливіших галузей економіки, що забезпечує продовольчу безпеку та стабільність суспільства. Одним із ключових факторів ефективного сільськогосподарського виробництва є точне прогнозування врожайності культур [11]. Прогнозування врожайності дозволяє фермерам та аграрним підприємствам приймати обґрунтовані рішення щодо планування посівів, управління ресурсами та підготовки до можливих коливань ринку.

В умовах сучасного розвитку технологій, інформаційні технології відіграють все більш важливу роль у сільському господарстві. Використання інформаційних технологій для прогнозування врожайності сільськогосподарських культур дозволяє суттєво підвищити точність прогнозів, оптимізувати процеси виробництва та знизити ризики, пов'язані з несприятливими погодними умовами, шкідниками та іншими факторами.

Традиційні методи прогнозування врожайності часто ґрунтуються на ручних спостереженнях та історичних даних, що може бути неточним та незручним інструментарієм [7]. Інформаційні технології, такі як штучний інтелект, машинне навчання та дистанційне зондування Землі, пропонують нові можливості для більш точного та своєчасного прогнозування.

Метою даної кваліфікаційної роботи є розробка та використання інформаційної технології для прогнозування врожайності сільськогосподарських культур. У процесі виконання роботи розглянуто сучасні методи та підходи до прогнозування врожайності, зокрема із застосуванням методів машинного навчання, аналізу великих даних та геоінформаційних систем (ГІС).

Робота містить теоретичний аналіз існуючих підходів до прогнозування врожайності, а також практичну частину, що включатиме розробку програмного забезпечення для збору, обробки та аналізу даних. Буде проведено експериментальну перевірку розробленої технології на основі реальних даних,

зібраних з різних джерел, включаючи супутникові знімки, метеорологічні дані та дані з сенсорів.

Очікується, що результати даної роботи зроблять внесок у підвищення ефективності сільськогосподарського виробництва, зменшення ризиків і втрат, пов'язаних з неврожайними роками, та покращення економічних показників аграрних підприємств.



## РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

### 1.1. Особливості прогнозування урожайності сільськогосподарських культур

Прогнозування врожайності сільськогосподарських культур є складним та багатограним процесом, що вимагає врахування широкого спектра факторів (рис. 1.1).

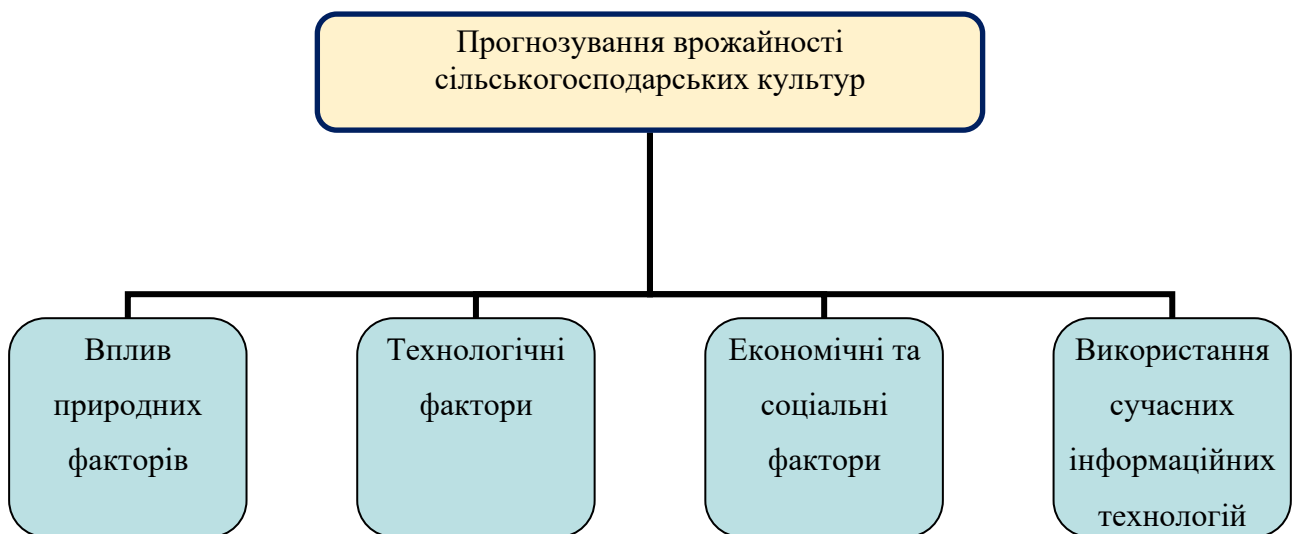


Рисунок 1.1 – Чинники, що впливають на результати прогнозування урожайності сільськогосподарських культур

Ефективне прогнозування врожайності дозволяє не лише оптимізувати агротехнічні заходи, але й приймати обґрунтовані економічні рішення, планувати використання ресурсів та знижувати ризики, пов'язані з несприятливими умовами.

Розглянемо особливості впливу природних факторів на процес прогнозування урожайності сільськогосподарських культур. Прогнозування врожайності значною мірою залежить від природних умов, таких як клімат, погодні умови, тип ґрунту та наявність водних ресурсів. Зміни клімату можуть суттєво впливати на вегетаційний період та стан рослин, що вимагає точних і своєчасних даних для коректного прогнозування.

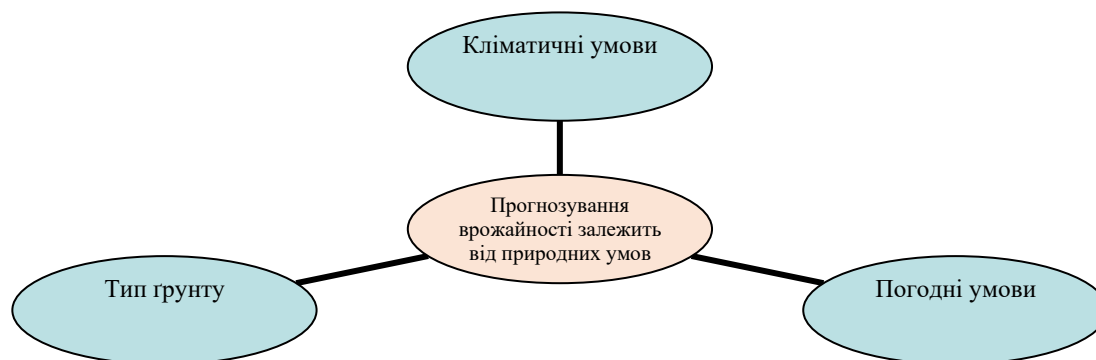


Рисунок 1.2 – Прогнозування врожайності сільськогосподарських культур залежить від природних умов

Кліматичні умови (температура, кількість опадів, тривалість сонячного освітлення та інші кліматичні показники) мають прямий вплив на ріст і розвиток рослин. Непередбачувані кліматичні зміни можуть суттєво вплинути на врожайність, тому важливо використовувати довгострокові метеорологічні прогнози та історичні дані.

Погодні умови (щоденні погодні зміни, такі як перепади температур, наявність опадів чи посухи) можуть істотно впливати на стан посівів. Використання сучасних метеорологічних моделей та сенсорних систем для моніторингу погодних умов є критичним для точного прогнозування.

Тип ґрунту (родючість, структура, текстура та інші фізико-хімічні властивості ґрунту) визначає можливість вирощування певних культур. Аналіз ґрунту та його відповідність вимогам конкретних культур є необхідним етапом у процесі прогнозування врожайності.

Сучасні агротехнології та методи обробки даних надають нові можливості для прогнозування врожайності (рис. 1.3).

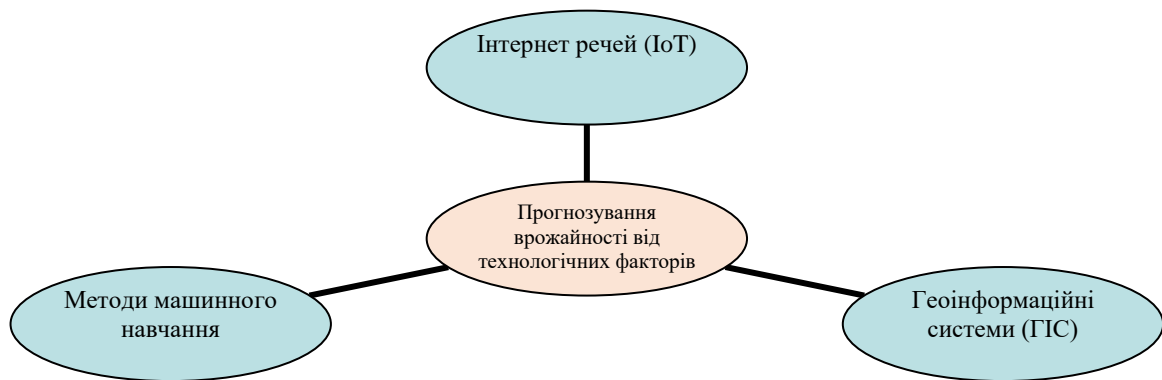


Рисунок 1.3 – Прогнозування врожайності сільськогосподарських культур залежить від технологічних складових

Використання алгоритмів машинного навчання дозволяє обробляти великі обсяги даних та виявляти складні закономірності між різними факторами, що впливають на врожайність. Це підвищує точність прогнозів та дає змогу адаптувати агротехнічні заходи до конкретних умов.

Геоінформаційні системи ГІС дозволяють інтегрувати просторові дані, такі як супутникові знімки, з іншими видами інформації (кліматичними, ґрунтовими тощо). Це забезпечує більш детальний аналіз та моделювання процесів, що відбуваються на полях.

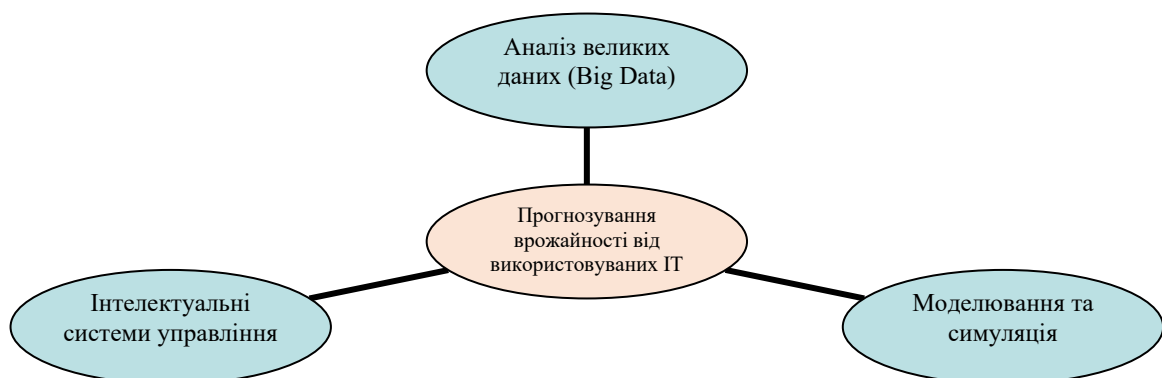


Рисунок 1.4 – Прогнозування врожайності сільськогосподарських культур залежить від використовуваних інформаційних технологій

Інтернет речей (IoT) передбачає використання сенсорних мереж для моніторингу стану ґрунту, рослин та погодних умов у режимі реального часу, що дозволяє оперативно реагувати на зміни та приймати обґрунтовані рішення щодо управління посівами.

Використання сучасних інформаційних технологій відкриває нові можливості для прогнозування врожайності. Однією із них є аналіз великих даних (Big Data). Збирання та аналіз великих обсягів даних з різних джерел дозволяє виявляти нові закономірності та робити більш точні прогнози.

Моделювання та симуляція відіграє важливу роль у прогнозуванні урожайності сільськогосподарських культур. Комп'ютерні моделі дозволяють проводити симуляції різних сценаріїв розвитку подій, оцінювати їх вплив на врожайність та розробляти оптимальні стратегії управління.

Інтеграція різних технологій в єдину систему управління дозволяє автоматизувати та інтелектуалізувати процеси збору даних, аналізу та прийняття рішень, що значно підвищує ефективність управління сільськогосподарськими процесами.

Отже, прогнозування врожайності сільськогосподарських культур є комплексним процесом, що вимагає врахування широкого спектра природних, технологічних, економічних та соціальних факторів. Використання сучасних інформаційних технологій дозволяє значно підвищити точність прогнозів та ефективність управління сільськогосподарськими процесами, що є ключовим фактором для забезпечення сталого розвитку аграрного сектора.

## **1.2. Оцінка врожайності сільськогосподарських культур за допомогою різних підходів**

Заслужовує на увагу використання алгоритмів машинного навчання для прогнозування урожайності сільськогосподарських культур. Пропонується використання на основі Sentinel-2 нормалізованого індексу різниці рослинності

(NDVI) як проксі для оцінки врожайності шляхом співвіднесення його з ССЕ піксельного рівня є звичайною практикою в сільському господарстві [25]. NDVI є надійним показником урожайності, оскільки він відображає здоров'я та ріст рослинності.



Рисунок 1.5 – Супутниковий моніторинг індексу NDVI у системі Soft.Farm [6]

Нормалізований індекс різниці рослинності (NDVI) кількісно визначає рослинність, вимірюючи різницю між ближнім інфрачервоним (яке рослинність сильно відбиває) і червоним світлом (яке рослинність поглинає)»:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1.1)$$

Отриманий NDVI на основі даних дистанційного зондування, він корелює з такими факторами, як вміст хлорофілу та індекс площі листя, надаючи розуміння розвитку врожаю (рис. 1.6).

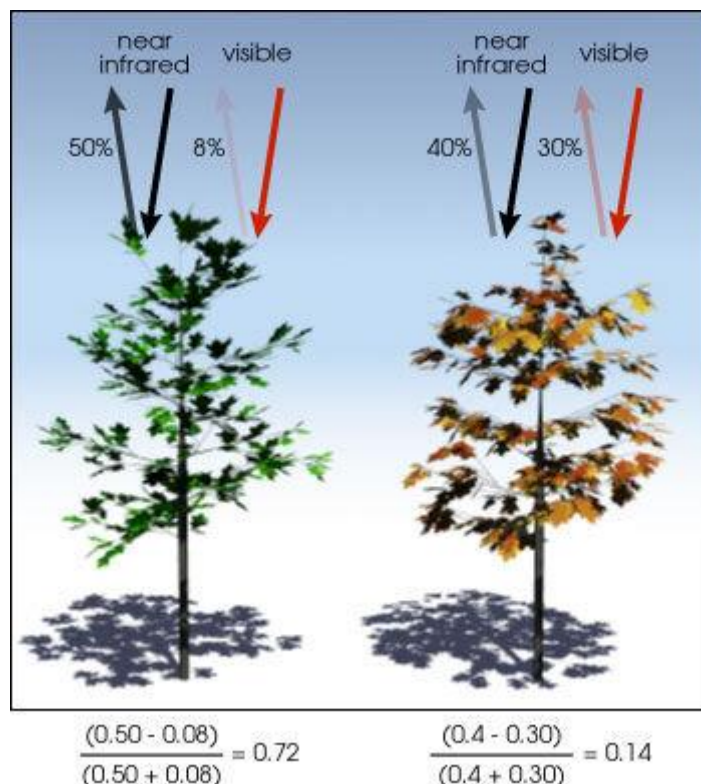


Рисунок 1.6 – Приклад визначення індексу NDVI у системі Soft.Farm [12]

Висока просторова та часова роздільна здатність роблять NDVI цінним інструментом для моніторингу великих сільськогосподарських територій. Допомогає виявити захворювання, дозволяючи вчасно втручатися. Однак для більш точних прогнозів важливо інтегрувати дані NDVI з іншою контекстною інформацією, такою як погодні умови та стан ґрунту.

Дослідження поєднує дані Sentinel-2 з даними наземних досліджень, щоб створити детальну та точну карту типів культур для сезону рабі (рис. 1.7). Методологія співвідносить базові сигнатури, отримані з наземних даних, із сигнатурами класів [15].

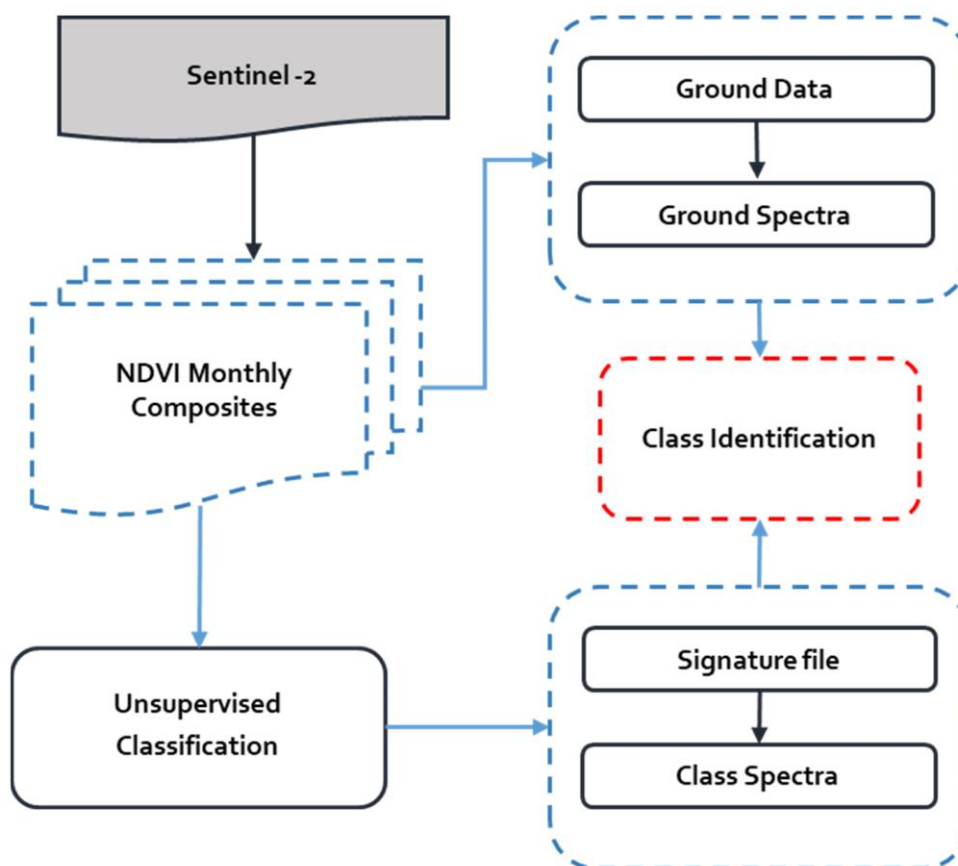


Рисунок 1.7 – Методологія, прийнята для картографування типів культур [15]

Для цього передбачаються фізичні вимірювання врожайності на певних ділянках поля. Ці експериментальні результати співвідносяться зі значеннями NDVI, отриманими із супутникових зображень у тих самих місцях. Також пропонується використовувати алгоритм машинного навчання випадкового лісу, використовуючи стек шарів NDVI як вхідні функції та навчальний набір даних. Цей підхід використовує алгоритм випадкового лісу для ефективного аналізу складних взаємозв'язків у складених шарах NDVI, що забезпечує точні прогнози на основі наданих навчальних даних.

Заслуговує на увагу імітаційна модель культури DSSAT. Модель CERES-Wheat, вбудована в структуру DSSAT v4.7.5, була використана для моделювання щоденного росту та розвитку посівів пшениці. Вхідні дані, що охоплюють параметри погоди, ґрунту, сорту/генотипу та врожаю, були зібрані з різних джерел, щоб забезпечити точність моделі (рис. 1.8).

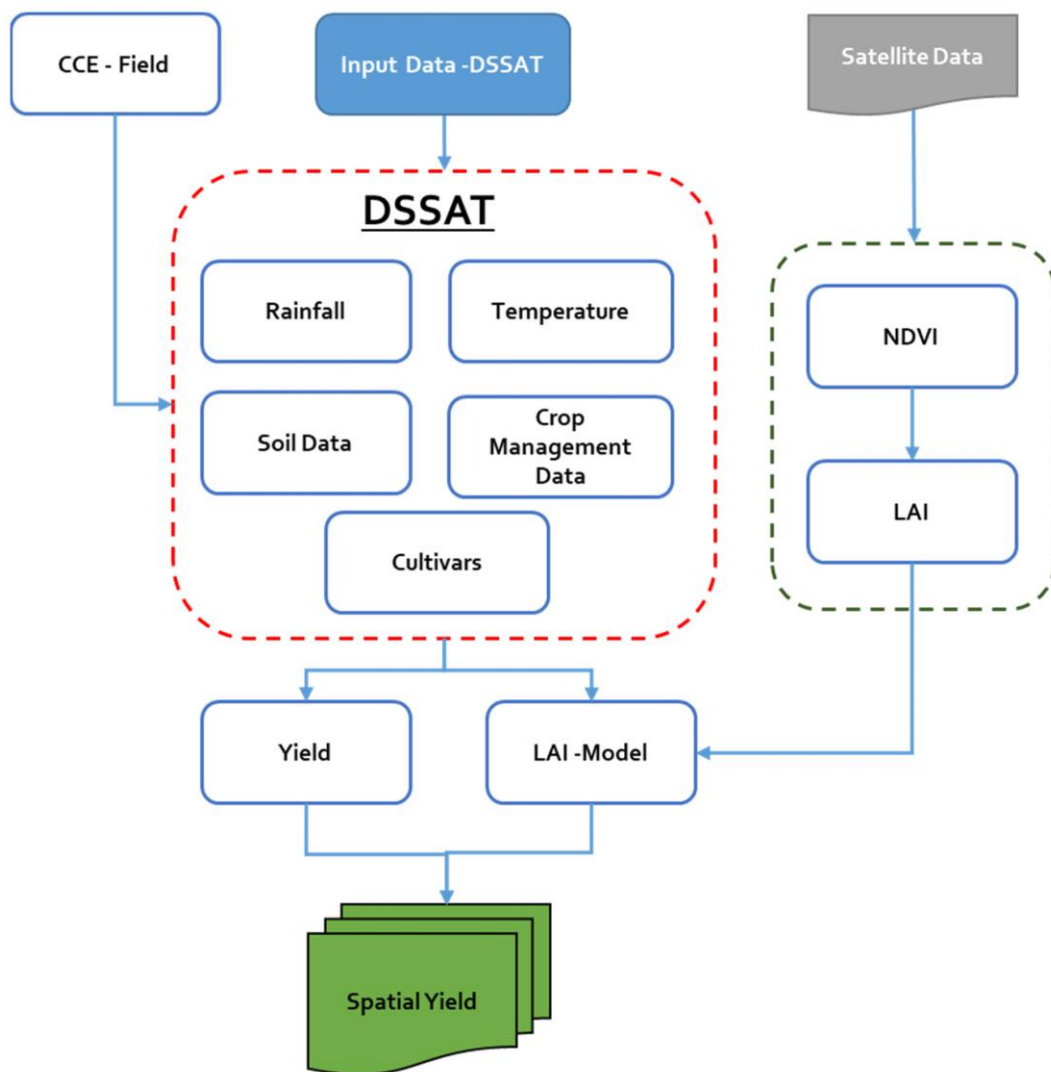


Рисунок 1.8 – Підхід до оцінки врожайності за допомогою DSSAT та її інтеграції з дистанційним зондуванням

Калібрування моделі було виконано з використанням даних, отриманих під час вегетаційного сезону пшениці рабі 2020–21 рр., і для визначення генетичних коефіцієнтів використовувався режим просторового аналізу. Відкалібрована модель пройшла перевірку шляхом порівняння її змодельованої врожайності з даними спостережень, що гарантує її надійність для майбутніх прогнозів. Щоб покращити інтеграцію результатів моделі з реальними умовами, значення індексу площі листя (LAI) були витягнуті із зображень NDVI.

Заслуговує на увагу напівфізичний підхід. Він являє собою комплексний напівфізичний метод для оцінки росту пшениці та оцінки врожайності, використовуючи багаточасові набори даних з різних джерел (рис. 1.9) [16].



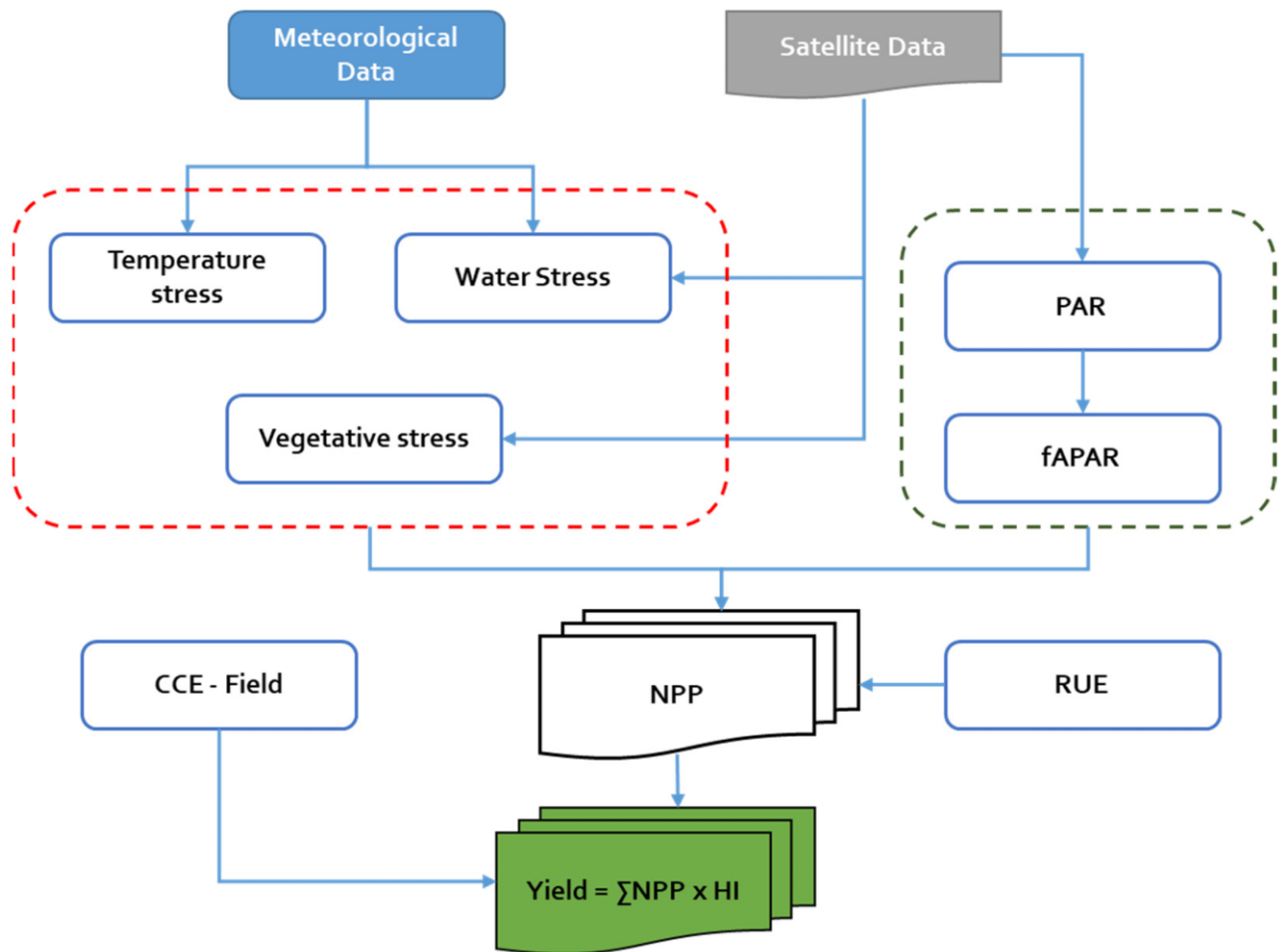


Рисунок 1.9 – Підхід до оцінки врожайності за допомогою SPM [16]

Використовуються дані класифікації сільськогосподарських культур Sentinel-2, щоденні дані про інсоляцію з просторовою роздільною здатністю 1 км, надані INSAT 3D, інформація MODIS про різні параметри та метеорологічні дані, надані NASA POWER, включаючи екстремальні температури.

Щоденні дані про інсоляцію з INSAT 3D, що становлять половину загальної інсоляції, були вирішальними. Крім того, використовувалися 8-денні композити fAPAR від MODIS з просторовою роздільною здатністю 500 м. Потім з 8-денним інтервалом обчислювали чистий первинний продукт (NPP) з використанням таких факторів, як PAR, fAPAR, стрес і максимальна ефективність використання радіації з просторовою роздільною здатністю 500 м.

$$NPP (g m^2 day1) = PAR \times fAPAR \times RUE \times stress. \quad (1.2)$$

Ефективність використання радіації (КЕ) пшениці коливається від 1,2 до 2,93 г МДж<sup>-1</sup>, тому ми прийняли загальне середнє значення КЕ як 2,06 г МДж<sup>-1</sup>.

Загальний NPP для всього сезону вирощування пшениці був отриманий із цих складених наборів даних. Урожайність пшениці була згодом розрахована шляхом поєднання загального NPP з індексом врожаю, отриманим від ССЕ.

### 1.3. Використання машинного навчання для прогнозування урожайності сільськогосподарських культур

Машинне навчання набуло популярності в сільському господарстві завдяки своєму успіху в інших галузях, таких як медицина, біоінформатика та обробка природної мови [22]. У роботах [24; 18] розглядають застосування машинного навчання в сільському господарстві.

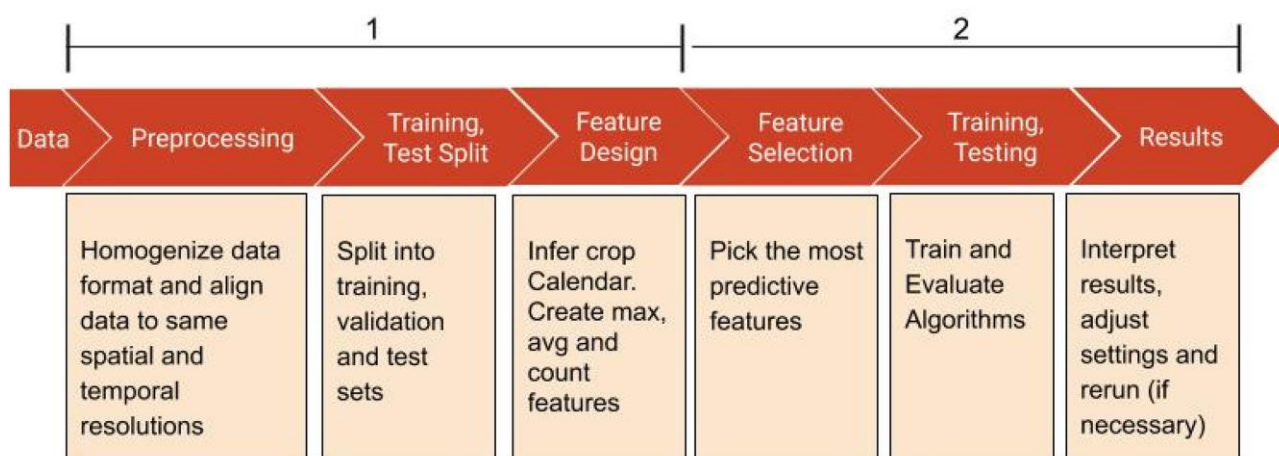


Рисунок 1.10 – Підхід машинного навчання до створення моделі прогнозування урожайності сільськогосподарських культур [24]

Багато досліджень (включених в огляди та інші) застосовували традиційне (або неглибоке) машинне навчання та глибоке навчання для прогнозування врожайності.

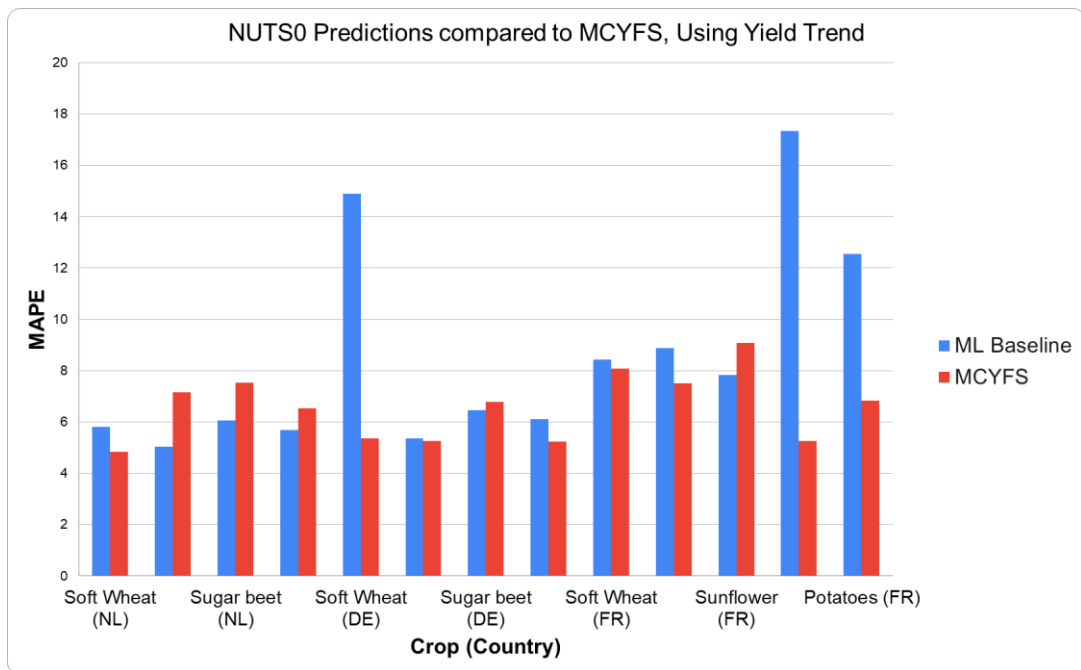


Рисунок 1.11 – Порівняння базового рівня машинного навчання із запропонованою моделлю прогнозування урожайності різних сільськогосподарських культур [24]

У роботі [18] порівняли чотири алгоритми ML як мета-моделі для прогнозування змодельованої врожайності кукурудзи та втрат азоту, навчених за доступними даними на момент посіву.

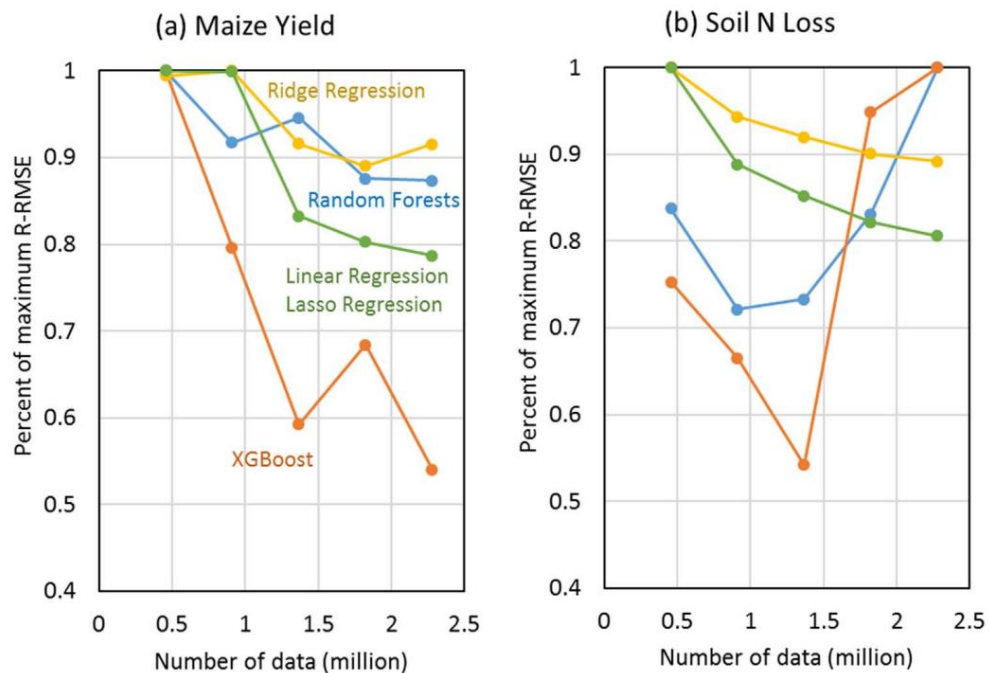


Рисунок 1.12 – Вплив зміни тривалості дослідження на прогнози врожайності кукурудзи (а) та втрати нітратів (б) (деякі лінії перекриваються) [18]

На основі отриманих результатів дійшли висновку, що прогнозована врожайність залежить від втрат азоту. Хоча більша продуктивність і менші вимоги до даних випадкових лісів роблять цей метод кращим серед мета-моделей, оцінених у цьому аналізі.

Рекомендується проводити оцінювання різних підходів ML для конкретних природно-виробничих умов із наявними наборами даних.

Серед застосувань неглибоких методів вчені побудували метамоделі машинного навчання на основі результатів моделі прогнозування врожаю APSIM.

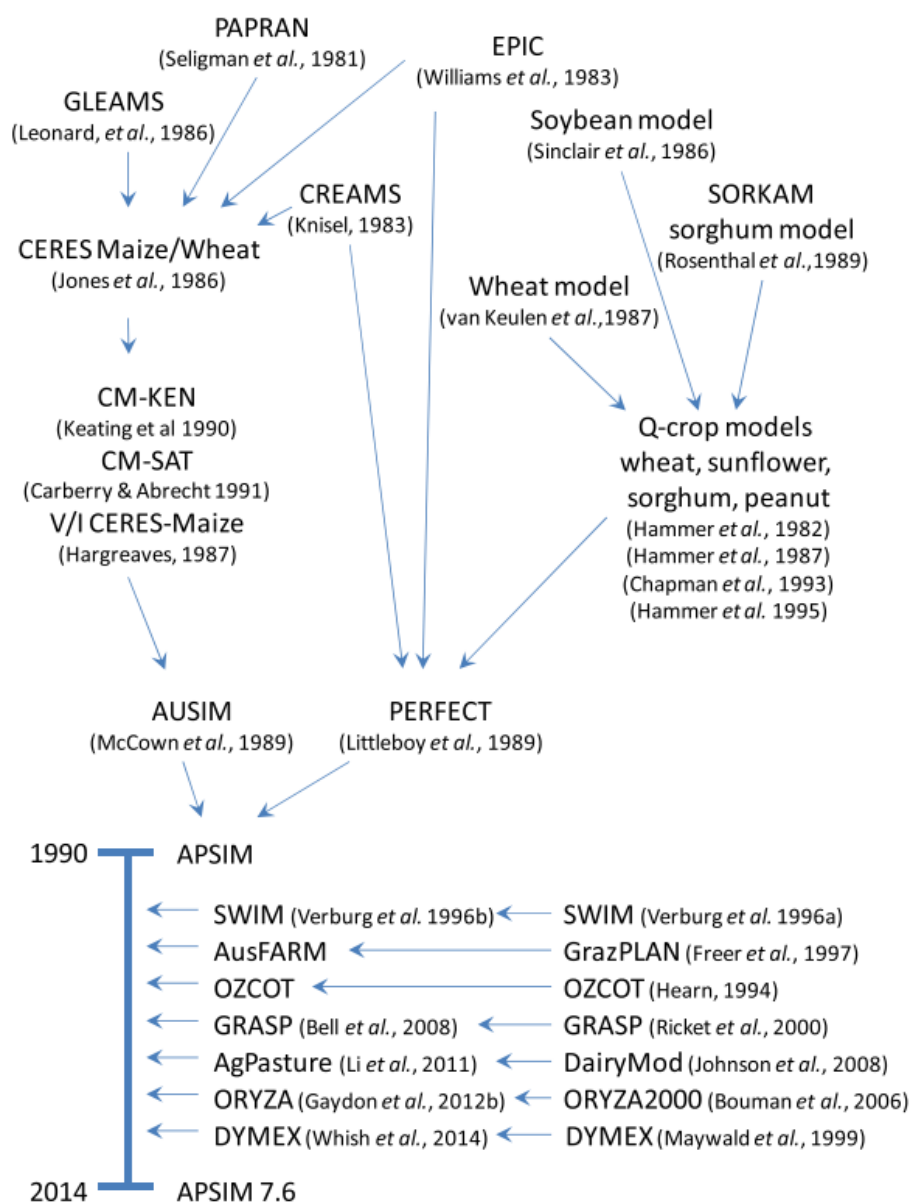


Рисунок 1.13 – Розвиток моделей APSIM (моделі, які вплинули на створення APSIM, та зовнішні моделі, які були включені до APSIM, після 1990)

Використання моделей за останнє десятиліття мало покращилося. Багато моделей прогнозування урожайності сільськогосподарських культур базуються на схожих підходах до ефективності використання сонячної радіації (RUE), але вони були реалізовані в різних базах вихідного коду або у принципово різні способи. Різні реалізації одного і того ж алгоритму можуть призводити до відмінностей, які не одразу помітні. Наприклад, DSSAT і APSIM використовують модель прогнозування урожайності, і деякі з їхніх параметрів виглядають схожими, але мають різні значення.

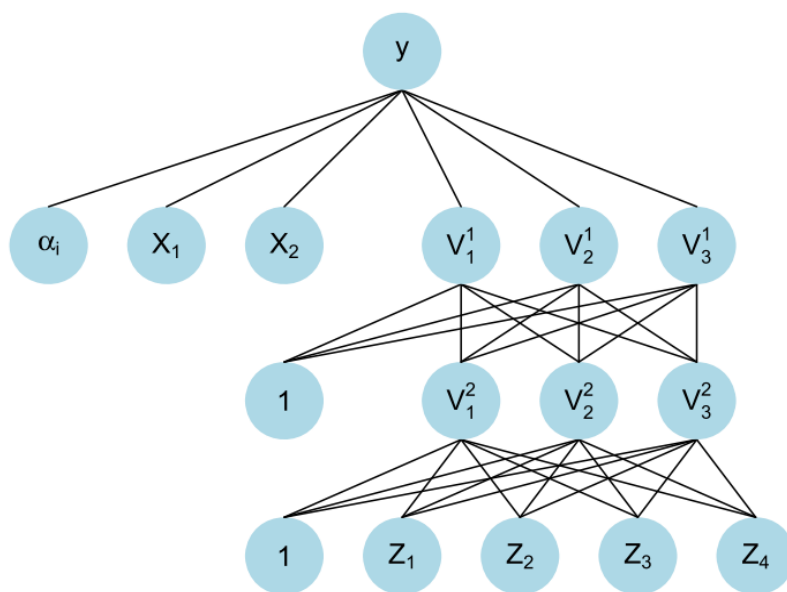


Рисунок 1.14 – Схематичне зображення (напівпараметричної) нейронної мережі [22]

Встановлено [22], що врожайність сільськогосподарських культур критично залежить від погоди. Запропоновано підхід до моделювання врожайності який використовує напівпараметричний варіант глибокої нейронної мережі, що може одночасно враховувати складні нелінійні зв'язки у високорозмірних наборах даних, а також відому параметричну структуру та неспостережувану поперечну неоднорідність. Використовуючи сценарії з набору кліматичних моделей, показали значний негативний вплив зміни клімату на врожайність кукурудзи. Дає набагато точніші результати ніж прогнозування урожайності за допомогою класичних статистичних методів.

#### 1.4. Сучасні інформаційні технології для прогнозування урожайності сільськогосподарських культур

У роботі представлена піраміда знань, що пов'язує дані з інформацією, знаннями і мудрістю, в якій дані є сировиною для розробки додатків, спрямованих на прийняття рішень (рис. 1.15).

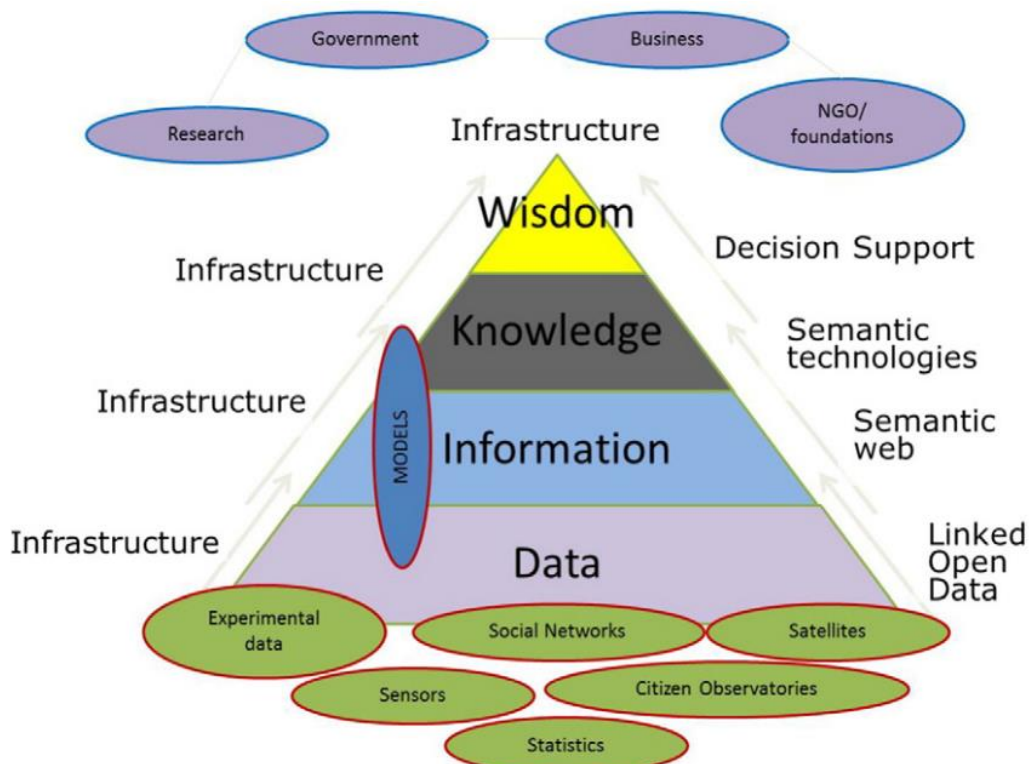


Рисунок 1.15 – Піраміда знань, що пов'язує дані з інформацією, знаннями і мудрістю [19]

Наступне покоління додатків, заснованих на моделюванні сільськогосподарських систем можуть допомогти компаніям, урядам та фермерам у харчовому ланцюгу приймати обґрунтовані рішення. Концепції ланцюгів знань (рис. 1.15) та ланцюгів застосування (рис. 1.16) надають взаємодоповнюючі перспективи на цінність і позиціонування моделювання в більш широкому контексті прийняття рішень та інформаційних технологій.

Ланцюг знань – це набір пов'язаних між собою кроків, за допомогою яких дані перетворюються на інформацію, знання і, зрештою, мудрість, яка

використовуються для прийняття рішень. Дані є сировиною, яка в поєднанні з описом і якісними характеристиками перетворюється на інформацію. Інформація може бути пов'язана з іншими джерелами інформації і розміщена в причинно-наслідкових ланцюжках для отримання знань.

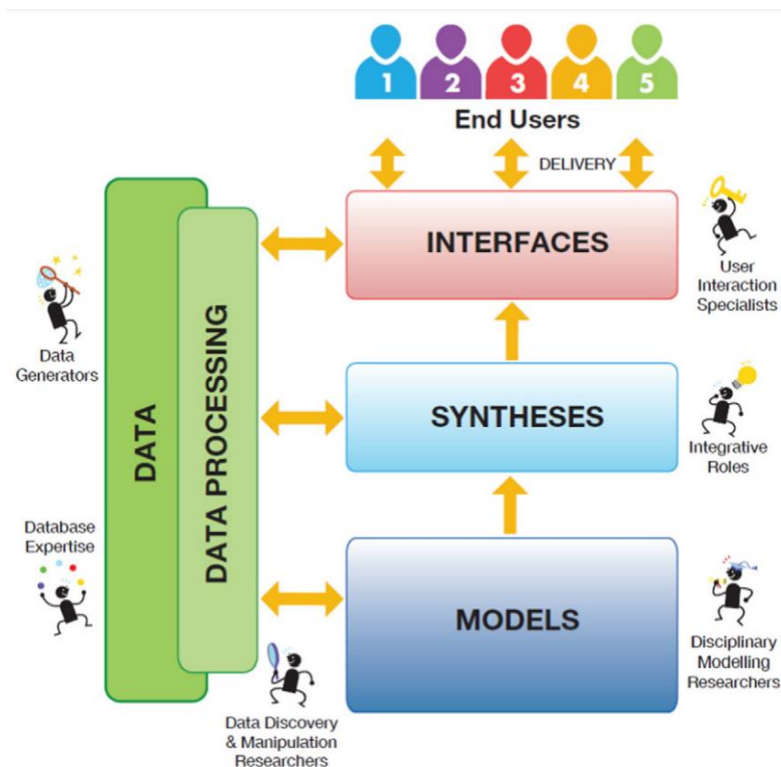


Рисунок 1.16 – Схема складових інформаційних технологій, що відображають потік даних та інформації через рівні моделювання, синтезу та інтерфейсу до кінцевих користувачів з роллю різних акторів

Оскільки сільськогосподарські підприємства відрізняються за розміром, наявністю робочої сили, ґрунтами та іншими характеристиками, запропоновано використовувати інструменти NextGen, щоб допомогти адаптувати поради для кожного господарства.

Отримується інформація від сільськогосподарських підприємств і вводиться у свій смартфон, на якому встановлені додатки NextGen, розроблені для фермерських систем регіону. Вони допомагають визначити комбінації компонентів системи, що можуть найкраще відповідати конкретним ситуаціям у господарстві. Це програмне забезпечення також надає шаблони файлів для інформаційних листів, написаних місцевою мовою (мовами), які описують

компоненти сільськогосподарських культур і систем землеробства, які, ймовірно, матимуть успіх у господарстві (рис. 1.17).

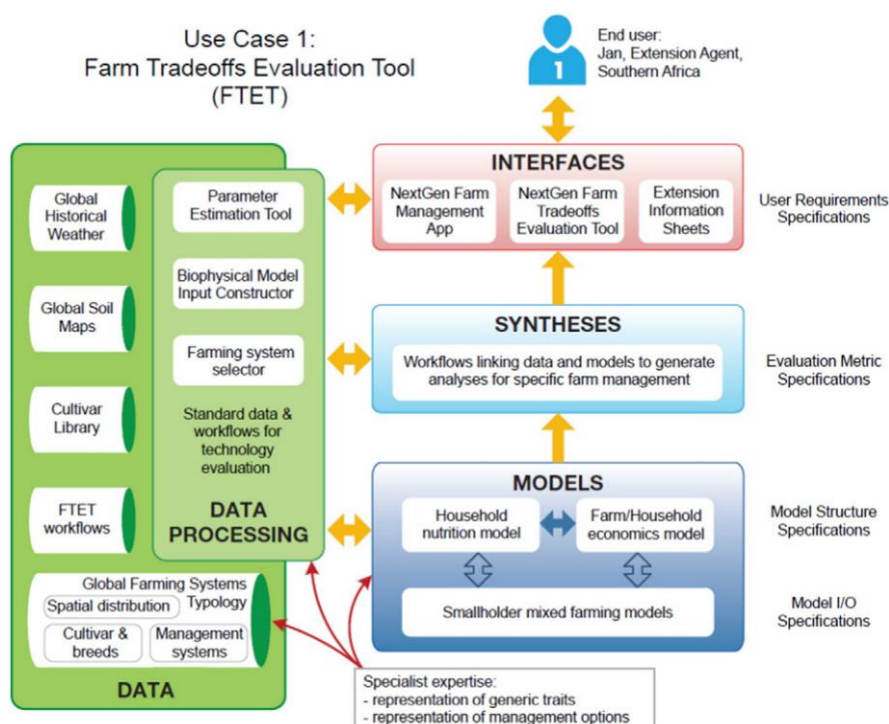


Рисунок 1.17 – Складові інформаційних технологій, що використовуються у сільськогосподарських підприємствах

Окремі технології мають обмежене практичне застосування для моделювання, оскільки кількісні дані не обробляються ефективно. Наприклад, онтології та тезауруси рідко включають специфікацію одиниць. Були зроблені зусилля, щоб заповнити цю прогалину [15]. На рис. 1.18 об'єднали словники з різних галузей сільського господарства, таким чином полегшивши зв'язок між моделями сільськогосподарських культур та економічними моделями.

Прогнозування врожайності сільськогосподарських культур є важливим завданням для забезпечення ефективного управління аграрним виробництвом. Сучасні інформаційні технології надають потужні інструменти для точного прогнозування врожайності, що дозволяє підвищити продуктивність та знизити ризики. Нами розглянуто основні технології, які використовуються для прогнозування врожайності сільськогосподарських культур.



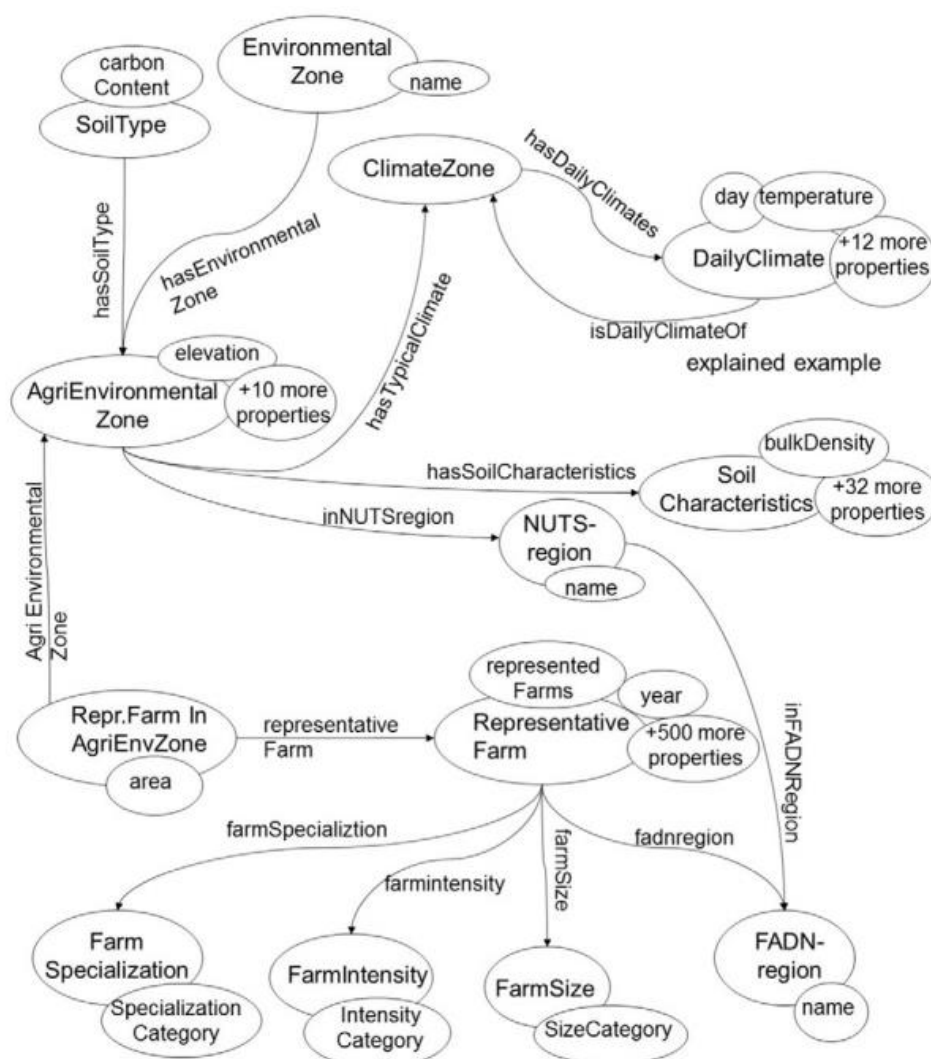


Рисунок 1.18. Діаграма «концепція-зв'язок», що відображає взаємозв'язок між фермерськими господарствами, інформацією про клімат та ґрунти [15]

Машинне навчання (ML) та штучний інтелект (AI) стали невід'ємними компонентами сучасних систем прогнозування врожайності. Вони дозволяють аналізувати великі обсяги даних та виявляти складні закономірності, що допомагають робити точні прогнози. Алгоритми машинного навчання, такі як лінійна регресія, рішення дерев та нейронні мережі, широко використовуються для прогнозування врожайності. Ці моделі можуть враховувати численні фактори, такі як кліматичні умови, тип ґрунту, агротехнічні заходи та інші, що впливають на врожайність. Глибокі нейронні мережі здатні обробляти великі масиви даних та виявляти приховані залежності між змінними. Вони ефективні

для аналізу супутникових знімків, метеорологічних даних та інших складних джерел інформації.

Геоінформаційні системи (ГІС) є потужним інструментом для збору, обробки та аналізу просторових даних, що мають важливе значення для прогнозування врожайності. Супутникові знімки дозволяють моніторити стан посівів, оцінювати площу посівів, щільність рослинного покриву та інші важливі показники. ГІС технології дозволяють інтегрувати ці дані з іншими джерелами інформації для точного прогнозування. Платформи ГІС дозволяють створювати детальні карти полів, моделювати просторовий розподіл врожайності та виявляти зони з різним рівнем продуктивності. Це допомагає у прийнятті рішень щодо оптимізації агротехнічних заходів.

Технології Інтернету речей (IoT) забезпечують постійний моніторинг умов вирощування сільськогосподарських культур у режимі реального часу. Різноманітні сенсори можуть вимірювати показники вологості ґрунту, температури, рівня освітленості, наявності поживних речовин та інших параметрів, що впливають на ріст рослин. Ці дані використовуються для оперативного коригування агротехнічних заходів. Інтеграція IoT з системами автоматичного управління дозволяє автоматизувати процеси зрошення, внесення добрив та інших агротехнічних операцій, що підвищує ефективність та точність управління.

Аналіз великих даних (Big Data) є важливим компонентом сучасних систем прогнозування врожайності, оскільки він дозволяє обробляти та аналізувати великі обсяги різноманітної інформації. Великі дані можуть включати метеорологічні дані, дані з сенсорів, супутникові знімки, історичні дані про врожайність та інші джерела інформації. Аналіз цих даних дозволяє виявляти тренди та закономірності, що впливають на врожайність. Використання алгоритмів аналізу великих даних дозволяє створювати складні прогностичні моделі, що враховують множину факторів та їх взаємодію. Це підвищує точність прогнозів та допомагає приймати обґрунтовані рішення.

Хмарні технології забезпечують доступ до обчислювальних ресурсів та інфраструктури, необхідної для обробки великих обсягів даних та виконання складних обчислень. Хмарні платформи дозволяють зберігати великі масиви даних та здійснювати їх обробку в реальному часі. Це забезпечує доступність даних та результатів аналізу для користувачів у будь-який час та з будь-якого місця. Хмарні технології дозволяють використовувати прогностичні моделі як послугу, що спрощує їх впровадження та використання аграрними підприємствами.

Дрони та безпілотні літальні апарати (БПЛА) є новітнім інструментом для збору даних про стан полів та посівів. Дрони можуть здійснювати високоточну аерофотозйомку та тепловізійний моніторинг полів, що дозволяє виявляти проблемні зони, оцінювати стан рослин та ефективність агротехнічних заходів. Використання дронів дозволяє здійснювати моніторинг стану посівів у реальному часі, що сприяє оперативному прийняттю рішень та коригуванню агротехнічних заходів.

Отже, сучасні інформаційні технології надають потужні інструменти для прогнозування врожайності сільськогосподарських культур. Використання машинного навчання, ГІС, IoT, аналізу великих даних, хмарних технологій та дронів дозволяє значно підвищити точність прогнозів та ефективність управління аграрним виробництвом. Це сприяє сталому розвитку сільського господарства та забезпеченню продовольчої безпеки.

### **1.5. Ідентифікація ідеї проекту**

Ідентифікація ідеї проекту є вагомим етапом у розробці ефективної інформаційної технології для прогнозування врожайності сільськогосподарських культур. Цей процес включає визначення основних цілей, завдань та компонентів проекту, а також аналіз потреб та вимог

користувачів. У цьому розділі буде розглянуто концепцію проекту, визначено його основні компоненти та обґрунтовано важливість реалізації.

Мета кваліфікаційної роботи полягає у розробка інформаційної технології для точного прогнозування врожайності сільськогосподарських культур з використанням сучасних методів аналізу даних та штучного інтелекту.

Отже, розробка інформаційної технології для точного прогнозування врожайності сільськогосподарських культур має велике значення для розвитку сучасного сільського господарства. Використання сучасних інформаційних технологій дозволить підвищити точність прогнозування врожайності, оптимізувати процеси та забезпечити стаке зростання аграрного сектора.

## РОЗДІЛ 2.

# ПРОЕКТУВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

### 2.1. Архітектура інформаційної технології прогнозування урожайності сільськогосподарських культур

Узагальнена структурна схема архітектури інформаційної технології прогнозування урожайності сільськогосподарських культур, яка передбачає 7 блоків, базується на моделі машинного навчання (навчених моделей RandomForestRegressor), представлена на рис. 2.1.

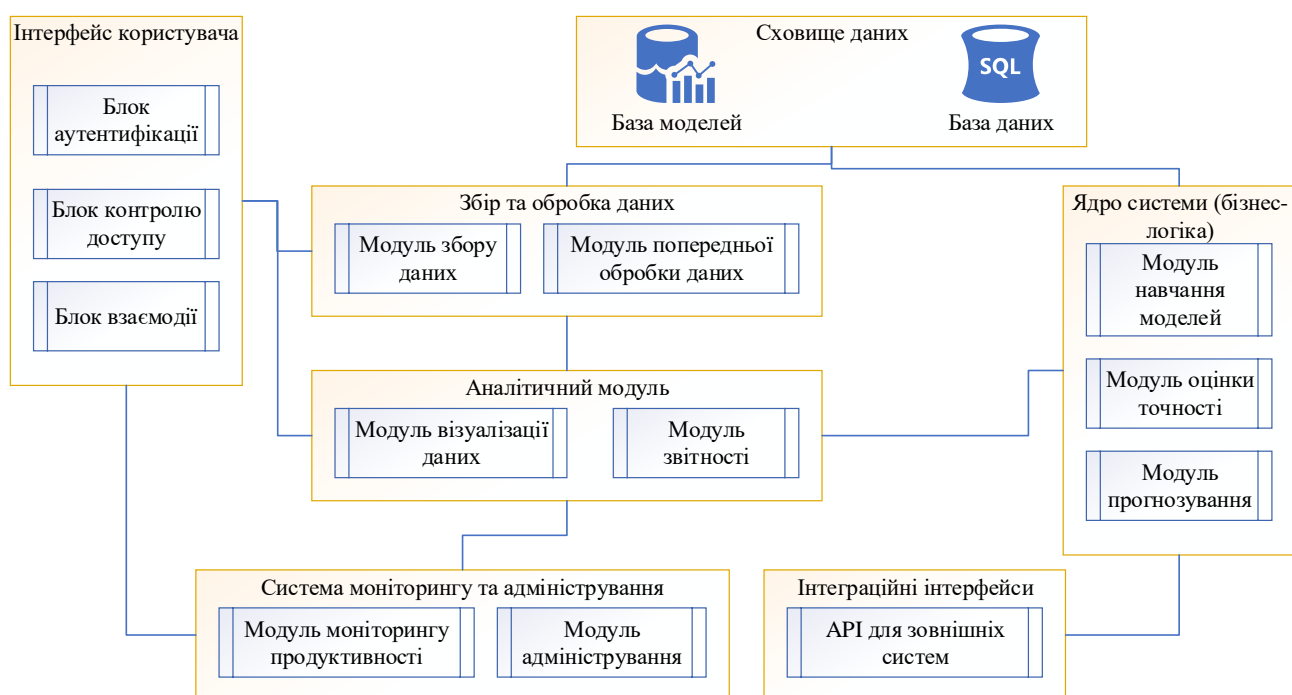


Рисунок 2.1 – Узагальнена структурна схема архітектури інформаційної технології прогнозування урожайності сільськогосподарських культур

#### *Інтерфейс користувача*

Інтерфейс користувача призначений для забезпечення взаємодії користувачів з інформаційною технологією прогнозування урожайності

сільськогосподарських культур. Основні функції інтерфейсу користувача включають:

Блок аутентифікації забезпечує вхід користувачів до системи з використанням імені користувача та паролю. Цей блок гарантує, що тільки авторизовані користувачі можуть отримати доступ до функцій системи. Він включає механізми реєстрації нових користувачів, відновлення забутих паролів та захист від несанкціонованого доступу.

Блок контролю доступу управляє рівнями доступу користувачів до різних функцій системи залежно від їхніх ролей. Це дозволяє розмежувати права та обов'язки користувачів, забезпечуючи безпеку та ефективність роботи системи. Наприклад, адміністраторам може бути надано доступ до всіх функцій, тоді як звичайні користувачі можуть мати доступ лише до обмеженого набору функцій, таких як перегляд прогнозів урожайності або генерація звітів.

#### *Збір та обробка даних*

Цей компонент інформаційної технології призначений для забезпечення якісного та своєчасного збору і обробки даних, необхідних для прогнозування урожайності сільськогосподарських культур. Основні функції включають:

Модуль збору даних відповідає за автоматизований збір даних з різних джерел, таких як метеорологічні служби, агрономічні бази даних, системи управління фермерськими господарствами та інші відповідні ресурси. Цей модуль забезпечує регулярне оновлення даних, необхідних для точного прогнозування. Він підтримує різні формати та протоколи передачі даних, включаючи API, FTP, веб-сервіси та інші методи інтеграції.

Модуль попередньої обробки даних виконує очищення, нормалізацію та перетворення зібраних даних у зручний для аналізу формат. Очищення включає видалення або виправлення помилок, відсутніх значень та аномалій. Нормалізація даних забезпечує приведення даних до єдиного масштабу або формату, що полегшує їх використання в моделі машинного навчання. Перетворення даних включає конвертацію даних у необхідні формати,

агрегування інформації та створення нових атрибутів на основі наявних даних, що підвищує точність та релевантність прогнозів.

### *База даних*

База даних є ключовим компонентом інформаційної технології прогнозування урожайності сільськогосподарських культур, забезпечуючи надійне зберігання та доступ до всіх необхідних даних і моделей. Основні складові бази даних включають:

Сховище даних забезпечує централізоване зберігання всіх зібраних даних у структурованому вигляді. Це включає як історичні, так і поточні дані, необхідні для навчання та роботи моделей прогнозування. Сховище даних організоване таким чином, щоб полегшити швидкий доступ до інформації, її аналіз та обробку. Структуровані дані можуть включати:

- ✓ історичні дані про урожайність сільськогосподарських культур;
- ✓ метеорологічні дані (середня кількість опадів, температура тощо);
- ✓ дані про використання пестицидів та інших агротехнічних заходів;
- ✓ інші релевантні показники, що впливають на урожайність.

База даних моделей призначена для зберігання навчених моделей машинного навчання, їхніх параметрів та метаданих. Це дозволяє легко керувати різними версіями моделей та їх налаштуваннями, забезпечуючи можливість оновлення та оптимізації прогнозів. Основні елементи бази даних моделей включають:

- ✓ збереження навчених моделей RandomForestRegressor та інших можливих моделей;
- ✓ параметри та гіперпараметри, використані під час навчання моделей;
- ✓ метадані, що описують процес навчання моделей, включаючи використані набори даних, обробку даних, результати валідації та тестування;
- ✓ логіни та історію змін моделей для забезпечення прозорості та відслідковування ефективності моделей у часі.

### *Ядро системи (бізнес-логіка)*

Ядро системи містить основну бізнес-логіку, необхідну для навчання моделей, прогнозування урожайності та оцінки точності прогнозів. Цей компонент відповідає за виконання критично важливих функцій системи прогнозування.

Модуль навчання моделей відповідає за навчання моделі `RandomForestRegressor` на основі зібраних та попередньо оброблених даних.

Основні функції цього модуля включають:

- ✓ завантаження даних із сховища даних;
- ✓ розділення даних на навчальний та тестовий набори;
- ✓ підбір оптимальних гіперпараметрів для моделі `RandomForestRegressor`;
- ✓ навчання моделі на навчальному наборі даних;
- ✓ збереження навчених моделей та їх параметрів у базі даних моделей.

Модуль прогнозування використовує навченої моделі для прогнозування урожайності на основі введених користувачем або нових зібраних даних.

Основні функції цього модуля включають:

- ✓ завантаження необхідних даних для прогнозування;
- ✓ використання навченої моделі для генерації прогнозів урожайності.
- ✓ обробка та форматування результатів прогнозування для подальшого використання та візуалізації.

Модуль оцінки точності відповідає за перевірку та оцінку точності прогнозів, які були згенеровані моделлю. Основні функції цього модуля включають:

- ✓ порівняння прогнозованих значень із фактичними даними;
- ✓ використання різних метрик оцінки точності, таких як середньоквадратична помилка (`mse`), середня абсолютна помилка (`mae`) та інші;
- ✓ аналіз результатів оцінки для виявлення можливих недоліків у моделі;



✓ корекція моделі на основі нових даних та результатів оцінки, включаючи повторне навчання моделі при необхідності.

Ці модулі взаємодіють між собою для забезпечення точного та надійного прогнозування урожайності сільськогосподарських культур, а також для постійного вдосконалення моделей на основі нових даних та результатів оцінки.

#### *Аналітичний модуль*

Аналітичний модуль призначений для обробки та представлення результатів прогнозування у зрозумілому для користувачів форматі. Він включає наступні ключові компоненти.

Модуль візуалізації даних відповідає за відображення результатів прогнозування у вигляді графіків, діаграм та інших візуальних елементів. Основні функції цього модуля включають:

✓ створення інтерактивних графіків та діаграм, які відображають результати прогнозування урожайності за різні періоди.

✓ візуалізація тенденцій та змін у даних, що допомагає користувачам зрозуміти динаміку урожайності.

✓ порівняння прогнозованих та фактичних значень для оцінки точності моделі.

✓ налаштування візуальних елементів для різних категорій користувачів, включаючи фермерів, агрономів та аналітиків.

Модуль звітності відповідає за генерацію звітів, що містять інформацію про прогнозовану урожайність, тенденції та рекомендації для користувачів. Основні функції цього модуля включають:

✓ автоматичне створення звітів на основі прогнозованих даних та результатів оцінки точності;

✓ інтеграція текстових та візуальних елементів (графіків, діаграм) у звіти для кращого розуміння інформації;

✓ генерація різних типів звітів (щоденних, щомісячних, річних) відповідно до потреб користувачів;

- ✓ надання рекомендацій щодо оптимізації агротехнічних заходів на основі прогнозованих даних;
- ✓ можливість експорту звітів у різні формати (pdf, excel, word) для зручного обміну та архівування.

Аналітичний модуль допомагає користувачам не тільки отримувати точні прогнози урожайності, але й легко інтерпретувати та використовувати ці дані для прийняття обґрунтованих рішень у сільськогосподарському виробництві.

#### *Система моніторингу та адміністрування*

Цей компонент системи відповідає за нагляд за роботою системи, відстеження її ефективності, а також за управління адміністративними аспектами, такими як налаштування системи та управління користувачами.

Модуль моніторингу продуктивності відстежує ефективність роботи системи та її компонентів. Основні функції цього модуля включають:

- ✓ моніторинг ресурсів системи, таких як використання процесора, обсяг пам'яті та мережевий трафік;
- ✓ відстеження часу відгуку системи на запити користувачів;
- ✓ аналіз прогресу навчання моделей та їх продуктивності;
- ✓ виявлення та вирішення можливих проблем та відхилень у роботі системи.

Модуль адміністрування відповідає за управління користувачами, налаштуваннями системи та регулярне оновлення даних. Основні функції цього модуля включають:

- ✓ управління доступом користувачів до різних функцій системи;
- ✓ налаштування параметрів системи, таких як періодичність оновлення даних, розмір зразка для навчання моделей та інші;
- ✓ реєстрація нових користувачів та управління їхніми обліковими записами;
- ✓ моніторинг та керування резервними копіями даних для забезпечення безпеки та надійності.

Система моніторингу та адміністрування забезпечує надійну та ефективну роботу системи прогнозування урожайності сільськогосподарських культур, забезпечуючи високу продуктивність та безпеку даних.

### *Інтеграційні інтерфейси*

Інтеграційні інтерфейси забезпечують можливість взаємодії між системою прогнозування урожайності та зовнішніми інформаційними системами та джерелами даних. Одним з ключових інтеграційних інтерфейсів є API (інтерфейс програмування додатків) для зовнішніх систем, який дозволяє системі обмінюватися інформацією з іншими програмами або платформами. Основні характеристики цього інтерфейсу включають API для зовнішніх систем. API для зовнішніх систем забезпечує інтеграцію з іншими інформаційними системами та джерелами даних для обміну інформацією. Основні функції цього API включають:

- ✓ можливість надсилання запитів на отримання даних з системи прогнозування урожайності, таких як прогнозовані значення урожайності, статистика моделей тощо;
- ✓ можливість надсилання даних для аналізу або використання в системі прогнозування урожайності, наприклад, історичні дані про урожайність, метеорологічні дані, дані про використання пестицидів тощо;
- ✓ забезпечення безпеки та аутентифікації для захисту обміну даними між системами;
- ✓ документація, що пояснює доступні ендпоінти, формати даних та методи запитів для інтеграції з іншими системами.

Цей API дозволяє легко і ефективно інтегрувати систему прогнозування урожайності з іншими системами та джерелами даних, що дозволяє розширити її можливості та використовувати більше інформації для точніших прогнозів.

## 2.2. Аналіз вимог до інформаційної технології

Нами виконано детальний аналіз вимог до інформаційної технології прогнозування урожайності сільськогосподарських культур. Аналіз включає в себе розгляд потреб користувачів, визначення функціональних та нефункціональних вимог до системи, а також формулювання основних функцій та характеристик системи (рис. 2.2).

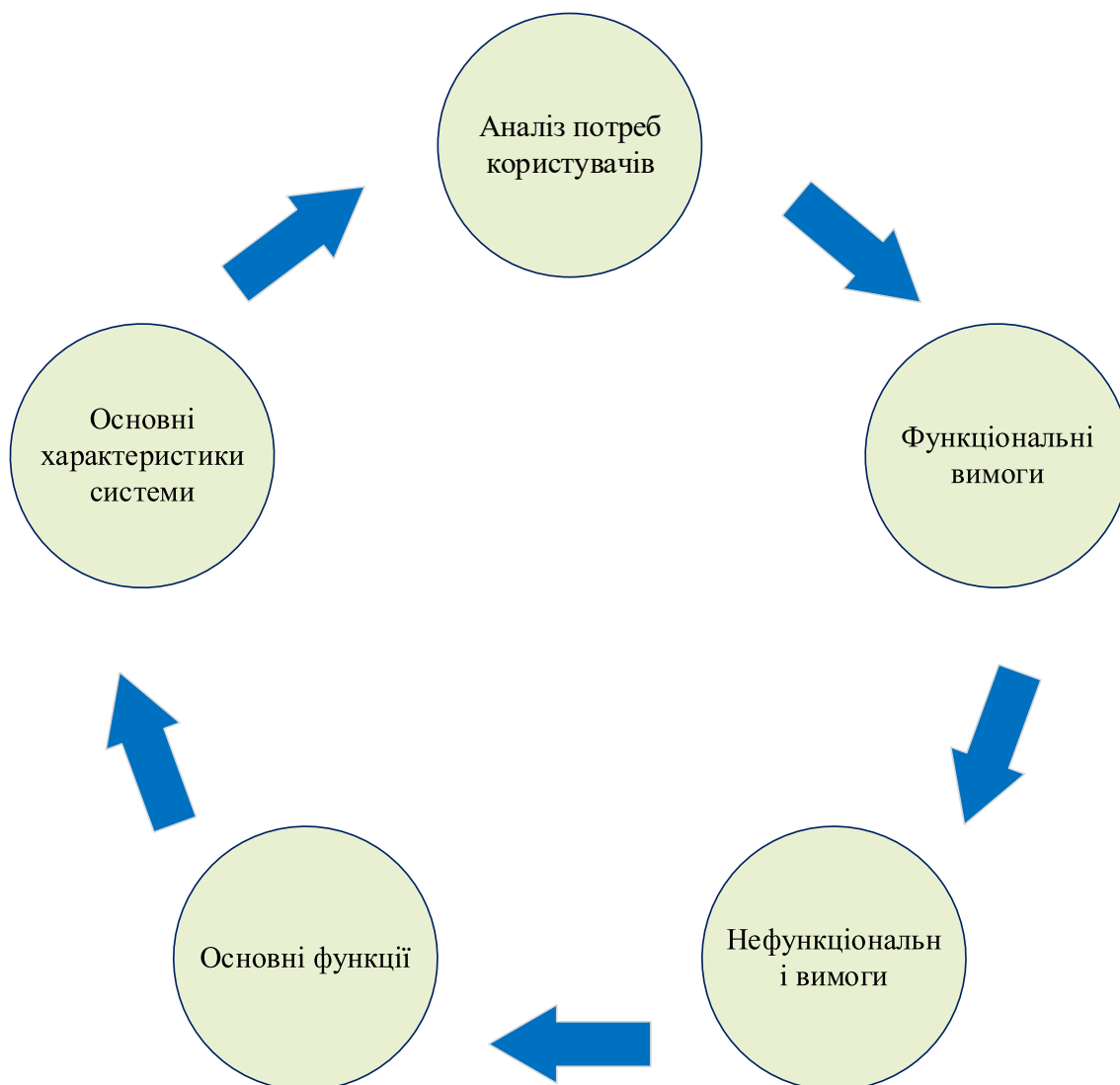


Рисунок 2.2 – Вимоги до інформаційної технології

### *Аналіз потреб користувачів*

*Сільські виробники.* Користувачі хочуть мати доступ до точних та надійних прогнозів щодо урожайності своїх культур. Вони цікавляться не лише

прогнозом загального врожаю, але і факторами, що впливають на нього, такими як погода, використання пестицидів та інші агротехнічні параметри.

*Агрономи.* Фахівці в галузі агрономії зацікавлені в інструментах для аналізу та впливу на урожайність сільськогосподарських культур. Вони потребують доступу до даних про погодні умови, ґрунтову вологу, використання добрив та інших факторів, щоб надати рекомендації фермерам щодо оптимальних методів обробки землі.

*Дослідники.* Вчені та дослідники в області сільськогосподарських наук зацікавлені в доступі до великих обсягів даних для аналізу та розробки нових методів прогнозування урожайності.

#### *Функціональні вимоги*

Система повинна забезпечувати можливість автоматичного збору та обробки різноманітних даних, включаючи дані про погоду, ґрунт, використання пестицидів тощо.

Система повинна використовувати алгоритми машинного навчання для побудови моделей прогнозування урожайності на основі зібраних даних.

Користувачам повинна бути надана можливість візуалізувати результати прогнозування у вигляді графіків та діаграм, а також генерувати звіти з рекомендаціями.

#### *Нефункціональні вимоги*

Система повинна бути швидкою та ефективною, здатною обробляти великі обсяги даних у реальному часі.

Від системи очікується висока надійність та стійкість до помилок.

Система повинна забезпечувати захист конфіденційності та цілісності даних.

#### *Основні функції та характеристики системи*

*Збір та обробка даних.* Автоматичний збір та обробка даних з різних джерел.

*Моделювання.* Використання алгоритмів машинного навчання для побудови моделей прогнозування.

*Візуалізація.* Відображення результатів прогнозування у вигляді графіків та діаграм.

*Звітність.* Генерація звітів з рекомендаціями для користувачів.

Аналіз вимог дозволяє чітко визначити потреби користувачів та встановити основні функції та характеристики системи, необхідні для задоволення цих потреб.

### **2.3. Вибір засобів та інструментів**

Нами проведений аналіз та вибір засобів та інструментів, необхідних для реалізації інформаційної технології прогнозування урожайності (рис. 2.3). Це включатиме вибір програмних мов, баз даних, бібліотек машинного навчання та інших інструментів.

Проведений аналіз і вибір технологій, необхідних для реалізації інформаційної технології прогнозування урожайності сільськогосподарських культур, дозволяє забезпечити ефективну, надійну та масштабовану систему. Нами вибрано мови програмування, бази даних, бібліотеки машинного навчання та інші інструменти, які будуть використовуватися для реалізації проекту.

Насамперед здійснено вибір мови програмування – Python. Це основна мова програмування для розробки системи прогнозування урожайності. Python вибрано за його потужні можливості в області обробки даних і машинного навчання, а також за наявність великої кількості бібліотек та інструментів. Основною перевагою цієї мови є простота у використанні, велика кількість бібліотек для аналізу даних та машинного навчання (numpy, pandas, scikit-learn). Вона вміщує бібліотеки NumPy, Pandas, Scikit-learn та Matplotlib для візуалізації. Причому SQL використовується для управління та взаємодії з реляційними базами даних. Перевагою є її поширення, стандартизація, ефективність у роботі з великими обсягами структурованих даних.

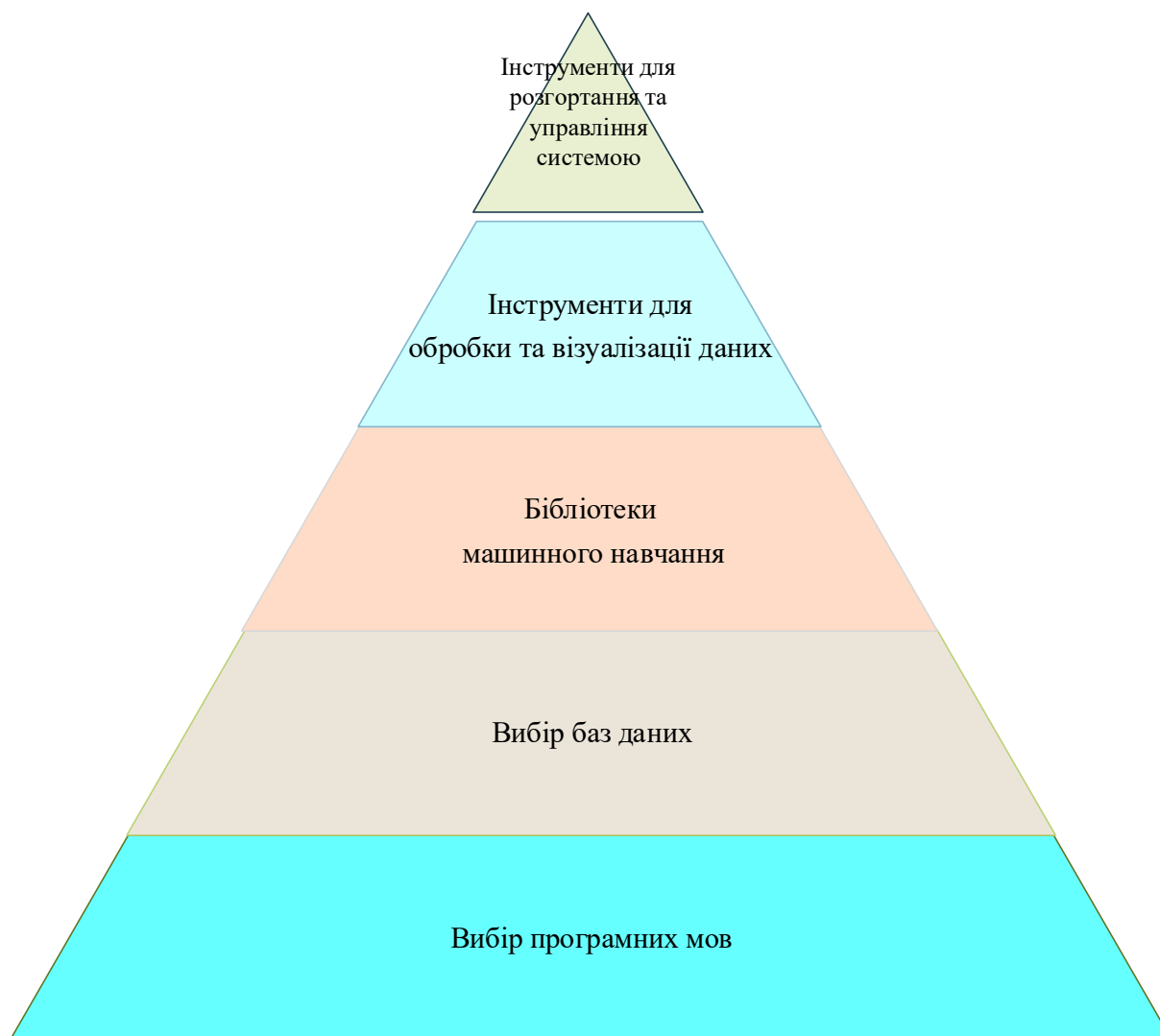


Рисунок 2.3 – Вибір засобів та інструментів

Щодо вибору баз даних, то пропонується PostgreSQL. Вибрана реляційна база даних забезпечує зберігання зібраних даних про погоду, урожайність, використання пестицидів тощо. Її перевагою є відкрите програмне забезпечення, підтримка складних запитів, висока продуктивність, підтримка великих обсягів даних. Також MongoDB використовується для зберігання неструктурованих даних та метаданих моделей машинного навчання. Перевагою є гнучкість у роботі з неструктурованими даними, горизонтальне масштабування, висока продуктивність при великих обсягах даних.

Нами вибрано бібліотеку машинного навчання – Scikit-learn. Це основна бібліотека для реалізації алгоритмів машинного навчання, включаючи RandomForestRegressor. Перевагою є простота у використанні, велика кількість

вбудованих алгоритмів, гарна документація. Також можливе використання TensorFlow/PyTorch для реалізації складніших моделей глибокого навчання, якщо це буде потрібно у майбутньому. Перевагою є можливості для розробки та навчання моделей глибокого навчання, підтримка GPU для пришвидшення обчислень.

Після цього нами вибрано інструменти для обробки та візуалізації даних. Запропоновано використовувати бібліотеку Pandas для обробки та аналізу даних. Її перевагою є простота у використанні, ефективність в роботі з табличними даними. Для візуалізації даних використовуються бібліотеку Matplotlib/Seaborn. Її перевагами є можливості для створення графіків та діаграм, легкість інтеграції з іншими бібліотеками Python.

Також вибрано інтерактивне середовище Jupyter Notebooks для розробки та тестування кодів. Перевагою є можливість інтерактивного аналізу та візуалізації даних, зручність для документування та демонстрації результатів.

Запропоновано використовувати інструменти для розгортання та управління системою – Docker. Він використовується для контейнеризації компонентів системи, що забезпечує їх незалежність і легкість розгортання. Перевагою є портативність, ізоляція середовищ, легкість масштабування. Також використовуватимемо Kubernetes – систему оркестрації контейнерів для управління розподіленими додатками. Її перевага – автоматизація розгортання, масштабування та управління контейнеризованими додатками. Окрім того вибрано Prometheus/Grafana – інструмент для моніторингу продуктивності системи та візуалізації метрик. Перевагою є достатні можливості для збору та аналізу метрик, інтерактивні інформаційні панелі.

Вибір засобів та інструментів здійснено з урахуванням потреб користувачів, вимог до системи та можливостей майбутнього розвитку. Python та бібліотеки машинного навчання, реляційні та не реляційні бази даних, інструменти для обробки та візуалізації даних, а також платформи для розгортання та моніторингу забезпечують надійну та ефективну реалізацію



інформаційної технології прогнозування урожайності сільськогосподарських культур.

## 2.4. Проектування інтерфейсу веб-додатку користувачів

Нами виконано проектування інтерфейсів користувача та інтеграційних інтерфейсів для взаємодії з системою прогнозування урожайності. Це включатиме розробку UI/UX дизайну, а також API для зовнішніх систем.

Інтерфейс користувача має бути інтуїтивно зрозумілим, зручним та забезпечувати доступ до всіх необхідних функцій системи (рис. 2.4). Основні функції інтерфейсу включають:

- ✓ введення даних для прогнозування;
- ✓ перегляд результатів прогнозування;
- ✓ аналіз даних за допомогою графіків та діаграм;
- ✓ генерація звітів.

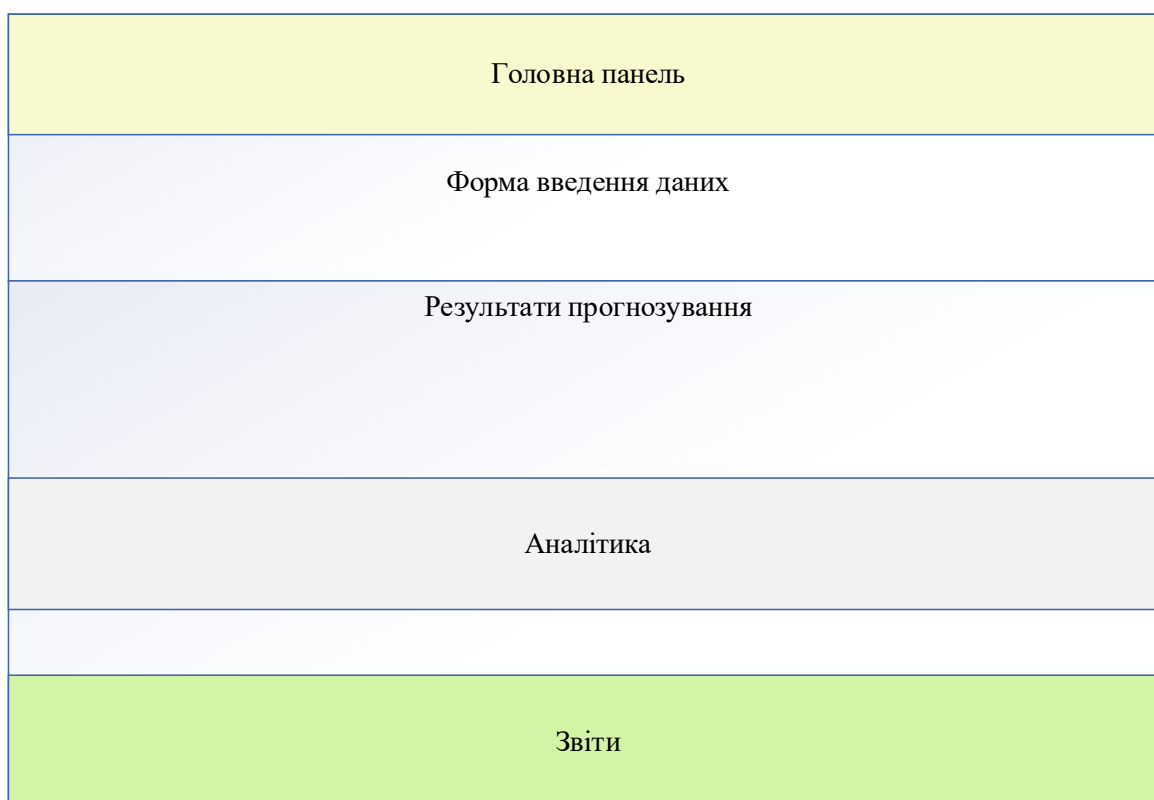


Рисунок 2.4 – Структура інтерфейсу користувачів

При цьому пропонується наступна структура інтерфейсу.

✓ *Головна панель.* Містить навігаційне меню з посиланнями на основні розділи системи: введення даних, результати прогнозування, аналіз, звіти.

✓ *Форма введення даних.* Інтерфейс для введення параметрів (область, рік, середня кількість опадів, температура тощо).

✓ *Результати прогнозування.* Відображає прогнозовану урожайність у вигляді таблиці та графіків.

✓ *Аналітика.* Інструменти для детального аналізу даних з візуалізацією.

✓ *Звіти.* Модуль для генерації та завантаження звітів у форматах PDF, Excel.

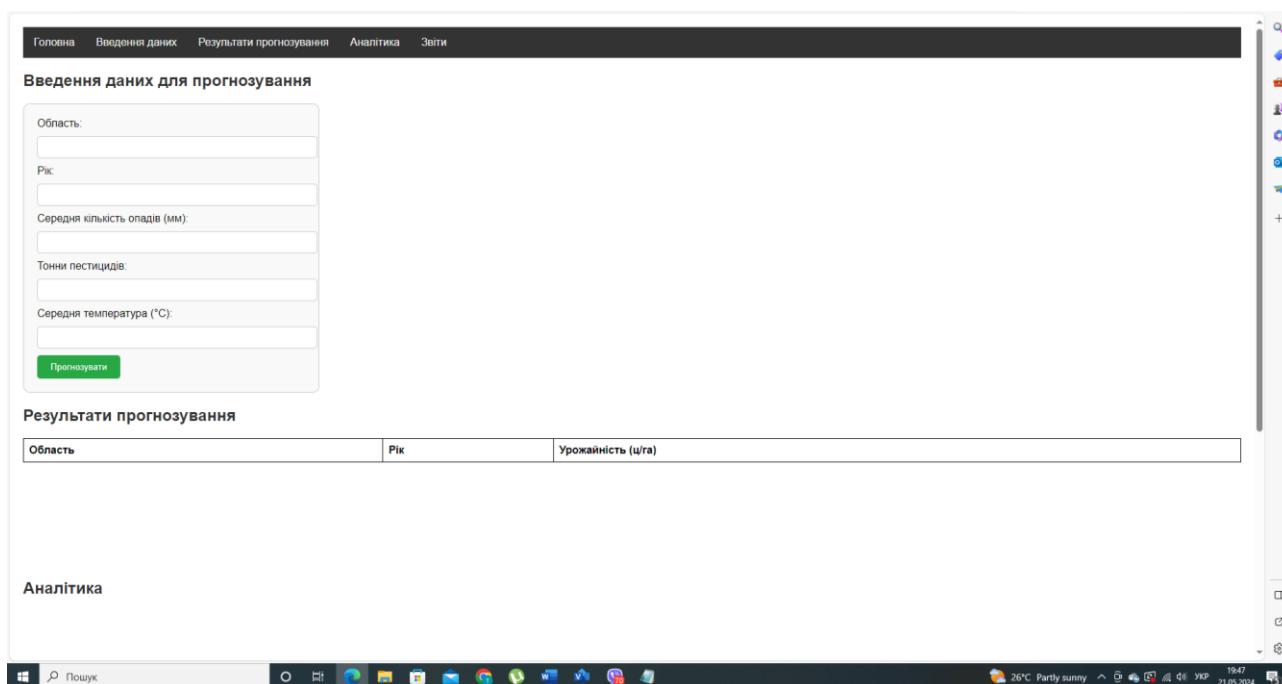


Рисунок 2.5 – Загальний вигляд інтерфейсу користувачів

Нами створено файл з розширенням .html (index.html) із HTML кодом форми введення даних (див. дод. А).

Також було створено файл styles.css для стилізації форми, що дало можливість зробити форму інтерфейсу користувачів більш привабливою. Створено файл script.js для обробки події відправки форми.

Запропонована структура інтерфейсу та написаний код дозволяють створити функціональний та зручний веб-додаток для прогнозування урожайності сільськогосподарських культур. Цей інтерфейс забезпечує користувачам можливість легко вводити необхідні дані, переглядати результати прогнозування та проводити аналітичний аналіз.

## 2.5. Інтеграційний інтерфейс (API)

API для зовнішніх систем забезпечує можливість інтеграції системи прогнозування з іншими інформаційними системами та джерелами даних. Основні функції API включають:

- ✓ отримання даних про погоду, сільськогосподарську культуру, використання пестицидів тощо;
- ✓ надсилання запитів для прогнозування урожайності;
- ✓ отримання результатів прогнозування та аналітики.

Структура API наступна:

- ✓ GET /api/v1/prediction – отримання результатів прогнозування на основі введених даних;
- ✓ POST /api/v1/data – введення нових даних для прогнозування;
- ✓ GET /api/v1/analysis – отримання аналітичних даних.

Нами написано код для реалізації API на Flask (Python). Цей код створює веб-додаток за допомогою фреймворку Flask для прогнозування урожайності сільськогосподарських культур на основі моделі машинного навчання, завантаженої з файлу. Нижче наведений детальний опис кожного фрагменту коду.

```

from flask import Flask, request, jsonify
import joblib
import pandas as pd

app = Flask(__name__)

# Завантаження навченої моделі з абсолютним шляхом
model = joblib.load('F:\\Dipl_2024\\Bak_2024\\It\\Slizep\\random_forest_model.pkl')

@app.route('/api/v1/prediction', methods=['POST'])
def predict():
    data = request.json
    input_data = pd.DataFrame([data])
    prediction = model.predict(input_data)
    return jsonify({'prediction': prediction[0]})

@app.route('/api/v1/data', methods=['POST'])
def add_data():
    data = request.json
    # Логіка збереження даних у базу даних
    return jsonify({'status': 'success'})

@app.route('/api/v1/analysis', methods=['GET'])
def analysis():
    # Логіка отримання аналітичних даних
    analysis_data = {'some_key': 'some_value'}
    return jsonify(analysis_data)

if __name__ == '__main__':
    app.run(debug=True)

```

Рисунок 2.6 – Код для реалізації API на Flask (Python)

Насамперед імпортуються необхідні бібліотеки – Flask, request, jsonify з пакету flask для створення веб-додатку та обробки HTTP-запитів. Joblib для завантаження моделі машинного навчання. Pandas для обробки даних у форматі DataFrame.

Після цього створюється екземпляр Flask-додатку. Відбувається завантаження навченої моделі. Завантажується навчена модель машинного навчання з файлу random\_forest\_model.pkl, що знаходиться за вказаним шляхом.

У подальшому визначається маршрут /api/v1/prediction, який обробляє POST-запити. Функція predict приймає JSON-дані з тіла запиту та перетворює ці дані у DataFrame.

Використовує завантажену модель для прогнозування на основі вхідних даних. Повертає результат прогнозування у форматі JSON.

Визначається маршрут /api/v1/data, який обробляє POST-запити.

Функція `add_data` приймає JSON-дані з тіла запиту. Повертає JSON-відповідь зі статусом успіху. Визначається маршрут `/api/v1/analysis`, який обробляє GET-запити. Повертає аналітичні дані у форматі JSON.

Запускає Flask-додаток у режимі налагодження. Це дозволяє автоматично перезапускати сервер при внесенні змін у код та забезпечує більш детальні повідомлення про помилки.

Отже, цей Flask-додаток забезпечує простий інтерфейс для взаємодії з моделлю машинного навчання для прогнозування урожайності. Він включає маршрути для отримання прогнозів, додавання даних та проведення аналітики.

### РОЗДІЛ 3.

## ЗБІР ТА ПОПЕРЕДНЯ ОБРОБКА ДАНИХ ДЛЯ НАВЧАННЯ МОДЕЛІ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

### 3.1. Набір даних для прогнозування урожайності сільськогосподарських культур

Нами розглядається набір даних, що використовується для прогнозування врожайності сільськогосподарських культур за допомогою алгоритмів машинного навчання (рис. 3.1). Набір даних включає різноманітні кліматичні змінні та інші фактори, що впливають на врожайність. У нашому наборі даних наявні наступні атрибути:

1. Area (Регіон) – географічний регіон, де вирощується сільськогосподарська культура;
2. Crop (Культура) – вид сільськогосподарської культури (наприклад, картопля, кукурудза, пшениця);
3. Year (Рік) – рік, за який зібрані дані;
4. Average rain fall mm per year (Середня кількість опадів на рік, мм) – середньорічна кількість опадів у міліметрах;
5. Pesticides tonnes (Пестициди, тон) – кількість використаних пестицидів у тонах;
6. Average temperature (Середня температура) – середньорічна температура у градусах Цельсія;
7. hg/ha\_yield (Врожайність, ц/га) – вихід продукції у центнерах на гектар (похідна змінна, що є метою прогнозування).

[69 rows x 7 columns]

Рядки для 'Area' = Ukraine:

	Area	Item	Year	average_rain_fall_mm_per_year	\
27357	Ukraine	Maize	2002	588	
27358	Ukraine	Potatoes	2002	588	
27362	Ukraine	Wheat	2002	588	
27363	Ukraine	Maize	2003	588	
27364	Ukraine	Potatoes	2003	588	
...	...	...	...	...	
27472	Ukraine	Potatoes	2022	588	
27476	Ukraine	Wheat	2022	588	
27477	Ukraine	Maize	2023	588	
27478	Ukraine	Potatoes	2023	588	
27482	Ukraine	Wheat	2023	588	
		pesticides_tonnes	avg_temp	hg/ha_yield	
27357		66764.00	9.24	26.229	
27358		66764.00	9.24	120.063	

Рисунок 3.1 – Фрагмент даних України про різноманітні кліматичні змінні та інші чинники, що впливають на врожайність

Для прогнозування врожайності сільськогосподарських культур буде використано вищезазначені змінні з використанням різних алгоритмів машинного навчання. Цей процес допомагає сільськогосподарським підприємствам та науковцям зрозуміти, як різні чинники, такі як кількість опадів, використання пестицидів та температура, впливають на врожайність різних культур. Точні прогнози врожайності є критично важливими для планування та управління аграрними процесами, а також для забезпечення продовольчої безпеки.

### 3.2. Підготовка даних для навчання моделі прогнозування урожайності сільськогосподарських культур

Насамперед нами виконано перевірку набору даних на наявність пропущених значень (рис. 3.2).

Заповнити пропущені значення середніми або медіанними значеннями відповідної змінної, або видалити записи з надто великою кількістю відсутніх даних.

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 12254 entries, 0 to 28241
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Area                                   12254 non-null  object
1   Item                                   12254 non-null  object
2   Year                                   12254 non-null  int64
3   average_rain_fall_mm_per_year        12254 non-null  int64
4   pesticides_tonnes                     12254 non-null  float64
5   avg_temp                              12254 non-null  float64
6   hg/ha_yield                           12254 non-null  float64
dtypes: float64(3), int64(2), object(2)
memory usage: 765.9+ KB

```

Рисунок 3.2 – Результати перевірки набору даних на наявність пропущених значень

Отриманий набір даних немає пропущених значень. У ньому є 12254 екземпляри даних, які характеризують є 3 унікальні значення – «Potatoes» (Картопля), «Maize» (Кукурудза) та «Wheat» (Пшениця). При чому їх кількість відповідно становить:

- ✓ Potatoes – 4276 (35%);
- ✓ Maize – 4121 (34%);
- ✓ Wheat – 3857 (31%).

Змінні, такі як «Area» (Регіон) та «Crop» (Культура), є категоріальними. Нами використано метод кодування, такий як label encoding, для перетворення категоріальних змінних у числовий формат, що прийнятний для моделей машинного навчання.

### 3.3. Результати аналізу впливу кліматичних умов на урожайність сільськогосподарських культур

Нами написано код, який дав можливість створити графіки впливу кількості опадів (рис. 3.3) та температури (рис. 3.5) на урожайність сільськогосподарських культур.



```

fig, axes = plt.subplots(1, 3, figsize=(15, 5))

# Plot for Potatoes
axes[0].plot(df0["average_rain_fall_mm_per_year"], df0["hg/ha_yield"], color='green', linestyle='-', marker='o')
axes[0].set_title('Potatoes')
axes[0].set_xlabel('Average Rainfall (mm/year)')
axes[0].set_ylabel('Yield (hg/ha)')
axes[0].grid(True)

# Plot for Maize
axes[1].plot(df1["average_rain_fall_mm_per_year"], df1["hg/ha_yield"], color='blue', linestyle='-', marker='o')
axes[1].set_title('Maize')
axes[1].set_xlabel('Average Rainfall (mm/year)')
axes[1].set_ylabel('Yield (hg/ha)')
axes[1].grid(True)

# Plot for Wheat
axes[2].plot(df2["average_rain_fall_mm_per_year"], df2["hg/ha_yield"], color='orange', linestyle='-', marker='o')
axes[2].set_title('Wheat')
axes[2].set_xlabel('Average Rainfall (mm/year)')
axes[2].set_ylabel('Yield (hg/ha)')
axes[2].grid(True)

plt.tight_layout()
plt.show()

```

Рисунок 3.3 – Код для створення графіків впливу кількості опадів на урожайність сільськогосподарських культур

У результаті отримано залежності урожайності сільськогосподарських культур від кількості опадів (рис. 3.4). Цей код дозволяє візуалізувати залежність врожайності трьох культур від середньої кількості опадів, використовуючи різні кольори та маркери для кожного виду культури.

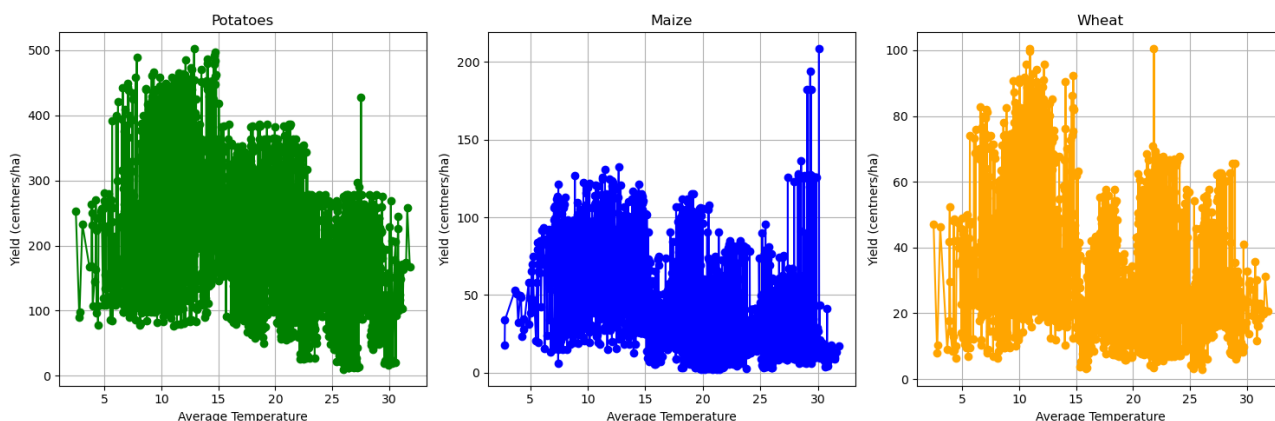


Рисунок 3.4 – Залежності урожайності сільськогосподарських культур від кількості опадів

Для трьох видів сільськогосподарських культур (картопля, кукурудза та пшениця) по осі X відображено середньорічну кількість опадів (у мм), а по осі Y - врожайність (ц/га).

Встановлено, що урожайності сільськогосподарських культур залежить від кількості опадів по різному для кожної із досліджуваних сільськогосподарських культур. При цьому за опадів більше 2000 мм / рік значно знижується врожайність. Найбільша врожайність спостерігається за 600...1800 мм/рік.

```
import seaborn as sns

# Створення фігури з сіткою 3x1
fig, axes = plt.subplots(3, 1, figsize=(10, 15))

# Встановлення заголовка фігури
fig.suptitle('Effect of Average Temperature on Yield', fontsize=16)

# Побудова точкових графіків з лінією тренду
sns.regplot(x=df0["avg_temp"], y=df0["hg/ha_yield"], ax=axes[0], color='red', line_kws={'color': 'blue', 'linewidth': 3})
axes[0].set_title('Potatoes')

sns.regplot(x=df1["avg_temp"], y=df1["hg/ha_yield"], ax=axes[1], color='blue', line_kws={'color': 'green', 'linewidth': 3})
axes[1].set_title('Maize')

sns.regplot(x=df2["avg_temp"], y=df2["hg/ha_yield"], ax=axes[2], color='green', line_kws={'color': 'orange', 'linewidth': 3})
axes[2].set_title('Wheat')

# Встановлення заголовків вісей
for ax in axes:
    ax.set_xlabel('Average Temperature')
    ax.set_ylabel('Yield (centners/ha)')
    ax.grid(True)

# Відображення графіків
plt.tight_layout()
plt.show()
```

Рисунок 3.5 – Код для створення графіків впливу середньорічної температури на урожайність сільськогосподарських культур

На основі запропонованого коду, який створює точкові графіки з лініями тренду для трьох видів сільськогосподарських культур (картоплі, кукурудзи та пшениці), можна зазначити наступне. Графіки чітко показують залежність врожайності від середньої температури для кожної культури. Це дозволяє візуально оцінити, як температура впливає на врожайність.

Лінії тренду, побудовані за допомогою `sns.regplot`, дають інтуїтивне розуміння загальної тенденції. Так як лінії мають нахил по спадаючій, то це свідчить що врожайність знижується з підвищенням температури.

Встановлено, що найвища урожайність картоплі припадає на середньорічну температуру 7...15 °С, водночас пшениці – 10...15 °С. За середньорічної температури 22 °С урожайність усіх культур значно знижується.

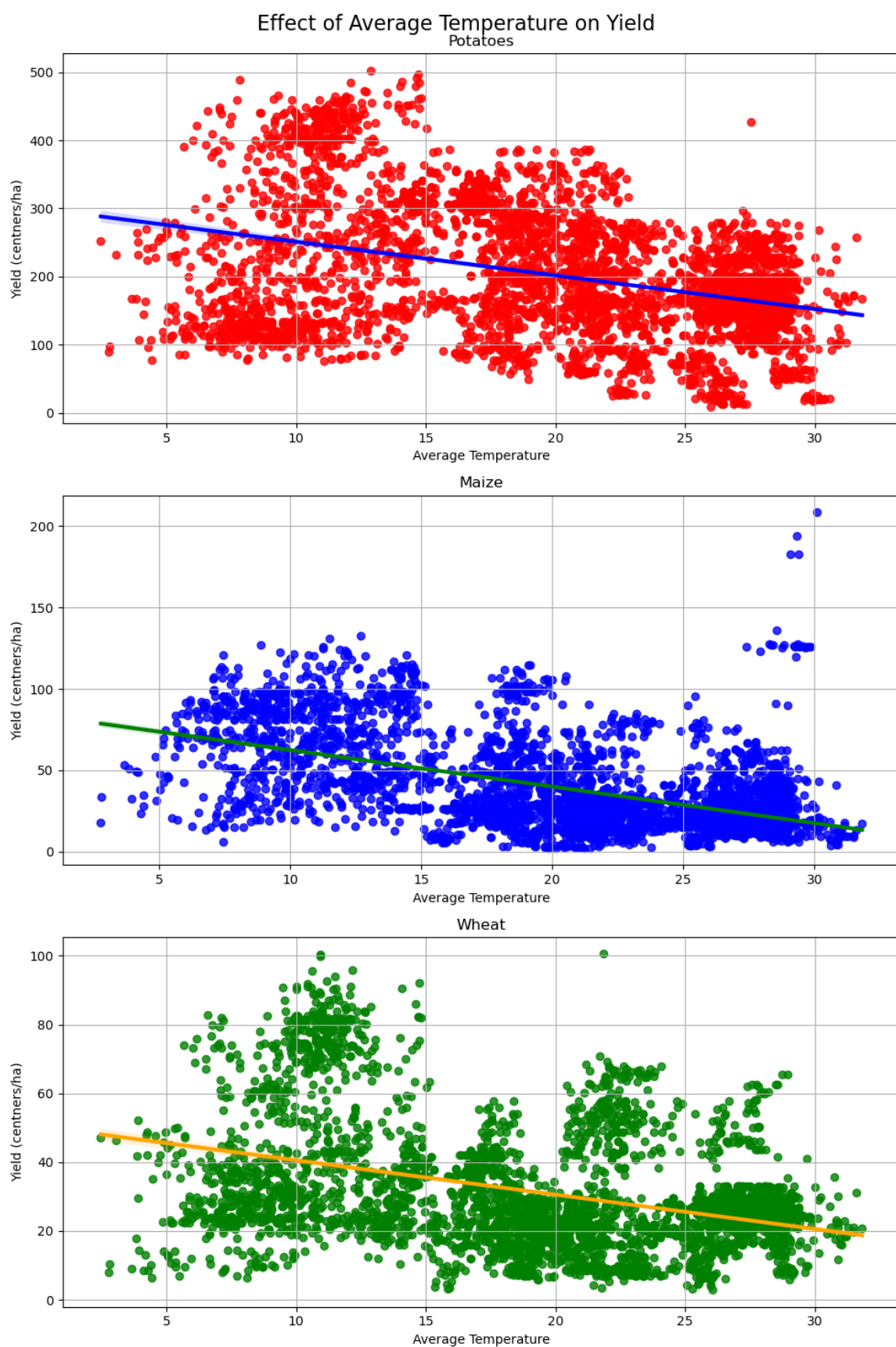


Рисунок 3.6 – Залежності урожайності сільськогосподарських культур від середньорічної температури у градусах Цельсія

## РОЗДІЛ 4.

### СТВОРЕННЯ ТА ВИКОРИСТАННЯ МОДЕЛІ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

#### 4.1. Кодування категоріальних змінних

Нами написано код (рис. 4.1), який використовує функцію `pd.get_dummies()` з бібліотеки `pandas` для кодування категоріальних змінних «Area» та «Item» у наборі даних `data`.

```
data = pd.get_dummies(data, columns=['Area', "Item"])
```

```
data.head(10)
```

Рисунок 4.1 – Код для кодування категоріальних змінних «Area» та «Item» у наборі даних `data`

Функція `pd.get_dummies()` перетворює кожну категорію в категоріальній змінній у новий стовпчик з двійковими значеннями (0 або 1) (рис. 4.2). Це робить змінні зручнішими для використання в машинному навчанні та інших статистичних методах.

	average_rain_fall_mm_per_year	pesticides_tonnes	avg_temp	hg/ha_yield	Area_Albania	Area_Algeria	Area_Angola	Area_Argentina	Area_Armenia	Area_Ai
0	1508	113.0	17.54	37.770	True	False	False	False	False	False
1	1508	113.0	17.54	67.824	True	False	False	False	False	False
5	1508	113.0	17.54	31.354	True	False	False	False	False	False
6	1508	113.0	16.53	30.225	True	False	False	False	False	False
7	1508	113.0	16.53	78.975	True	False	False	False	False	False
11	1508	113.0	16.53	21.855	True	False	False	False	False	False
12	1508	113.0	17.23	26.033	True	False	False	False	False	False
13	1508	113.0	17.23	84.077	True	False	False	False	False	False
17	1508	113.0	17.23	25.545	True	False	False	False	False	False
18	1508	113.0	17.22	25.342	True	False	False	False	False	False

Рисунок 4.2 – Фрагмент набору даних `data` із результатами кодування категоріальних змінних «Area» та «Item»

Для змінної `Area` функція `pd.get_dummies()` створено новий стовпчик для кожного унікального значення. Наприклад, якщо в `Area` є 3 унікальні значення (`A1`, `A2`, `A3`), то буде створено 3 нових стовпчика: `Area_A1`, `Area_A2`, `Area_A3`.

У кожному з цих нових стовпчиків буде значення 1 для рядків, де Area дорівнює відповідному унікальному значенню, і 0 для решти рядків.

Аналогічно, для змінної Item було створено нові стовпчики для кожного унікального значення в ній. У кожному з цих нових стовпчиків буде значення 1 для рядків, де Item дорівнює відповідному унікальному значенню, і 0 для решти рядків.

Кодування категоріальних змінних призводить до збільшення розміру набору даних. Важливо вибрати відповідну стратегію кодування для ваших даних та задачі.

Зокрема, для побудувати моделі машинного навчання, яка прогнозує врожайність сільськогосподарських культур на основі різних чинників, включаючи регіон, тип сільськогосподарської культури, використання пестицидів, кількість опадів та температура. Кодування змінних Area та Item за допомогою `pd.get_dummies()` допоможе вашій моделі краще зрозуміти ці категоріальні змінні та зробити більш точні прогнози.

## **4.2. Навчання моделей для прогнозування урожайності сільськогосподарських культур**

Прогнозування врожайності сільськогосподарських культур є важливою задачею, яка може допомогти сільськогосподарським підприємствам приймати кращі рішення щодо управління своєю діяльністю. Завдяки точному прогнозуванню врожайності можна оптимізувати використання ресурсів, таких добрива та пестициди, а також мінімізувати ризики втрат врожаю.

Існує багато різних алгоритмів машинного навчання, які можна використовувати для прогнозування врожайності сільськогосподарських культур.

Насамперед нами здійснено імпорт потрібних модулів та вибору моделей (рис. 4.3).

```

from sklearn.neural_network import MLPRegressor
from sklearn.svm import SVR
from sklearn.neighbors import KNeighborsRegressor
from sklearn.gaussian_process import GaussianProcessRegressor

# Список моделей для навчання, включаючи неймережеві моделі
models = {
    'RandomForestRegressor': RandomForestRegressor(),
    'BaggingRegressor': BaggingRegressor(),
    'GradientBoostingRegressor': GradientBoostingRegressor(),
    'AdaBoostRegressor': AdaBoostRegressor(),
    'DecisionTreeRegressor': DecisionTreeRegressor(),
    'LinearRegressor': LinearRegression()
}

```

Рисунок 4.3 – Код для імпорту потрібних модулів та вибору моделей

Після цього виконується створення словника моделей. Словник `models` дозволяє легко керувати кількома моделями, надаючи зручний спосіб доступу до них за допомогою рядкових ключів. Це спрощує процес навчання та оцінки різних моделей у циклі, після чого можна легко перераховувати моделі за їх іменами та застосовувати до кожної однакової процедури навчання та оцінки.

Після цього виконуємо поділ даних на тренувальні та тестові набори (рис. 4.4).

```

# Розділяємо дані на тренувальні та тестові
X = data.loc[:, data.columns != 'hg/ha_yield']
Y = data['hg/ha_yield']
x_train, x_test, y_train, y_test = train_test_split(X, Y, test_size=0.2, random_state=42)

```

Рисунок 4.4 – Код для поділу даних на тренувальні та тестові набори

Цей код розділяє дані на тренувальні та тестові набори, щоб підготувати їх для навчання та оцінки моделей машинного навчання.

Параметр `test_size=0.2` показує, що 20% даних буде використано для тестування, а 80% – для навчання. Параметр `random_state=42` встановлює фіксоване значення для генератора випадкових чисел, щоб забезпечити відтворюваність результатів. Це означає, що кожен раз, коли ви запускаєте цей код, розділення даних буде однаковим.

Наступним кроком є навчання моделей прогнозування урожайності сільськогосподарських культур (рис. 4.5).

```

import time
import pandas as pd
import numpy as np
from sklearn.metrics import r2_score, mean_squared_error

# Функція для обчислення скоригованого коефіцієнта детермінації
def adjusted_r2_score(r2, n, p):
    return 1 - (1 - r2) * (n - 1) / (n - p - 1)

# Ваш цикл навчання моделей
for model_name, model in models.items():
    start_time = time.time()
    model.fit(x_train, y_train)
    predictions = model.predict(x_test)
    r2 = r2_score(y_test, predictions)
    adj_r2 = adjusted_r2_score(r2, x_test.shape[0], x_test.shape[1])
    rmse = np.sqrt(mean_squared_error(y_test, predictions))
    time_taken = time.time() - start_time

    new_row = pd.DataFrame({
        'Model': [model_name],
        'Adjusted R-Squared': [adj_r2],
        'R-Squared': [r2],
        'RMSE': [rmse],
        'Time Taken': [time_taken]
    })

    results = pd.concat([results, new_row], ignore_index=True)

print(results)

```

Рисунок 4.5 – Код для навчання моделей прогнозування урожайності сільськогосподарських культур

Наведений код виконує навчання моделей машинного навчання, оцінює їхню ефективність за різними показниками та дає результат у DataFrame results. Для кожної моделі виконується навчання, використовуючи метод .fit() із вхідними даними поточного набору x\_train та відповідними мітками y\_train.

Після навчання моделей виконується прогнозування значень на тестовому наборі x\_test за допомогою методу predict().

Для оцінки ефективності моделей обчислюються такі метрики:

- ✓ коефіцієнт детермінації ( $R^2$ ) за допомогою функції r2\_score;
- ✓ скоригований визначення коефіцієнтів за допомогою функції adjusted\_r2\_score;
- ✓ середньоквадратичне відхилення (RMSE) за допомогою функції np.sqrt(mean\_squared\_error).

Після проходження всіх моделей виводяться результати - DataFrame results, який містить інформацію про кожну модель та її ефективність у прогнозуванні. Цей код дозволяє автоматизувати процес навчання та оцінки

кількох моделей машинного навчання та вивести результати в зручному табличному форматі.

Таблиця 4.1 – Результати оцінки ефективності моделей для прогнозування урожайності сільськогосподарських культур

№ п/п	Модель	Adjusted R-Squared	R-Squared	RMSE	Time Taken
1	RandomForestRegressor	0.981	0.982	13.142	5.939
2	BaggingRegressor	0.979	0.980	13.724	0.636
3	GradientBoostingRegressor	0.901	0.905	30.229	2.076
4	AdaBoostRegressor	0.813	0.822	41.461	1.011
5	DecisionTreeRegressor	0.974	0.975	15.409	0.097
6	LinearRegressor	0.812	0.820	41.655	0.036

Встановлено, що найкращою моделлю з цих результатів є RandomForestRegressor, оскільки вона має найвище значення скоригованого R-квадрату та найнижче значення середньоквадратичного відхилення (RMSE). RandomForestRegressor має трохи кращі показники за всіма цими показниками, що можна вважати її найкращою моделлю для прогнозування врожайності сільськогосподарських культур. Для RandomForestRegressor числові значення метрики ефективності такі. Скоригований R-квадрат становить 0,981, R-Squared (Коефіцієнт детермінації) – 0,982, RMSE (Середньоквадратичне відхилення) – 13.142 та тривалість навчання моделі – 5.939 секунд.

Ці значення свідчать про те, що модель RandomForestRegressor досить точно використовує урожайність сільськогосподарських культур на основі вхідних даних. Високі значення скоригованого та звичайного коефіцієнтів детермінації (близько до 0,98) свідчать про те, що більшість варіацій урожайності вдається пояснити моделі.

Отже, модель RandomForestRegressor може бути рекомендована для подальшого використання в прогнозуванні врожайності сільськогосподарських



культур, оскільки вона має найкращі показники ефективності серед інших представлених моделей.

### 4.3. Оцінення процесу навчання моделі RandomForestRegressor та точності прогнозування урожайності сільськогосподарських культур

Оцінення процесу навчання моделі RandomForestRegressor виконаємо на основі побудованого графіка, який відображає зміну точності моделі на збільшення розміру процесу навчального набору даних (рис. 4.6).

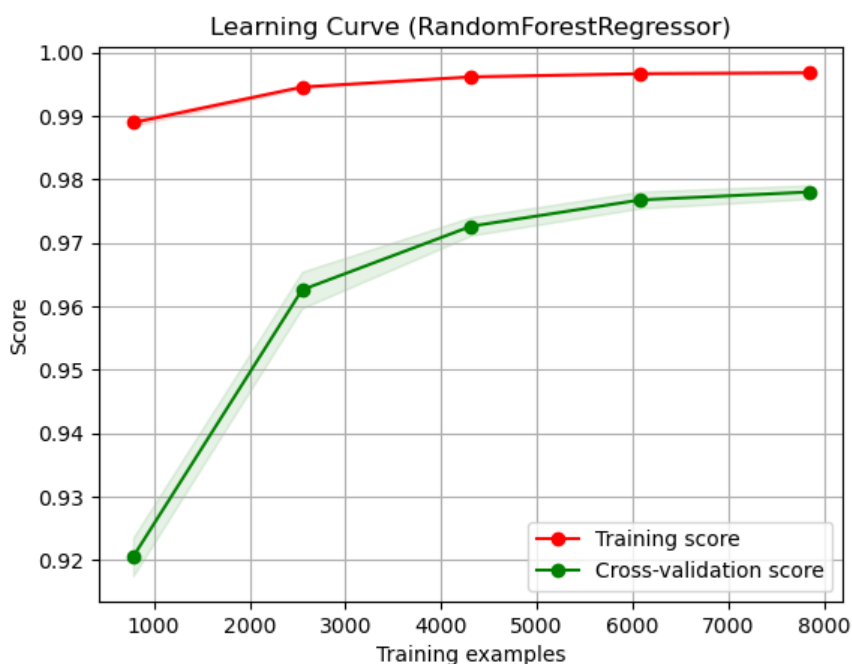


Рисунок 4.6 – Графік процесу навчання моделі RandomForestRegressor

Із графіка кривої навчання для моделі RandomForestRegressor видно, що починаючи з невеликого обсягу даних, модель має високу точність на навчальному наборі даних, після чого вона незначно зростає за збільшення обсягу даних. Проте точність на перехресній валідації спочатку невелика, однак зі навчального набору точність на перехресній валідації зростає, що показує, що модель краща узагальнюється на нових даних.

На останньому кроці виконано побудову графіка розсіювання для порівняння фактичних та прогнозованих значень урожайності (рис. 4.7). Кожна точка на графіку відображається фактичним значенням на осі X і відповідним прогнозованим значенням на осі Y.

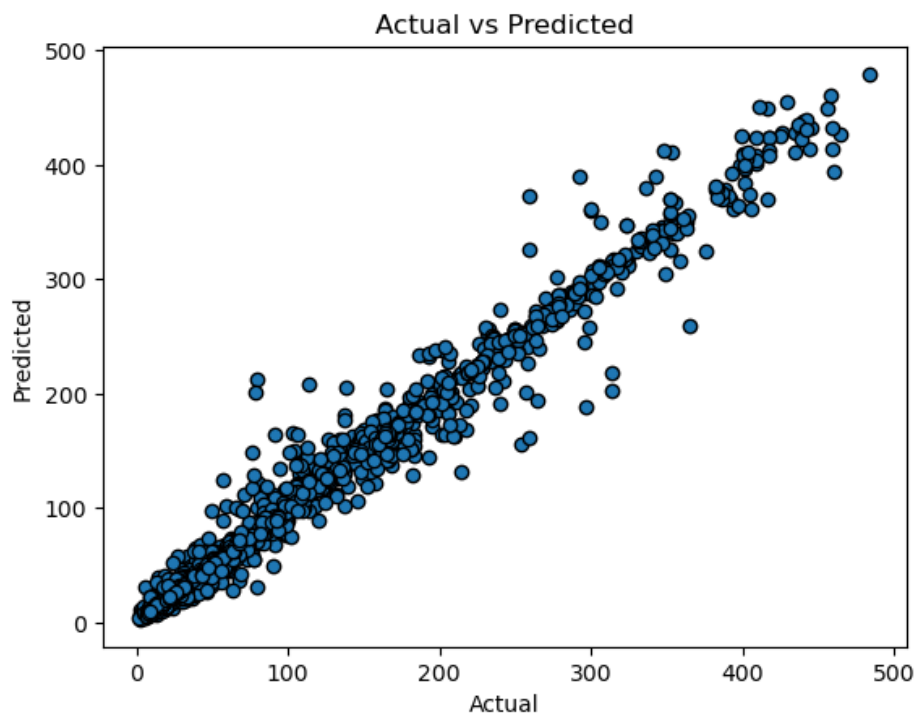


Рисунок 4.7 – Графіка розсіювання для порівняння фактичних та прогнозованих значень урожайності сільськогосподарських культур

За допомогою цього графіка бачимо, модель `RandomForestRegressor` добре прогнозує значення урожайності сільськогосподарських культур. Точки на графіку розсіювання розташовані близько до діагональної лінії, це означає, що модель `RandomForestRegressor` успішно прогнозує значення урожайності на основі вхідних даних.

## РОЗДІЛ 5. ОХОРОНА ПРАЦІ

Створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур пов'язане з певними ризиками для здоров'я та безпеки працівників [3]. Важливо вжити заходів для мінімізації цих ризиків та забезпечення безпечного та здорового робочого середовища.

Дотримання правил охорони праці при створенні та використанні інформаційних технологій прогнозування урожайності сільськогосподарських культур є запорукою безпечного та здорового робочого середовища, а також запобігання нещасних випадків та професійних захворювань.

### **5.1. Вимоги до мікроклімату у приміщенні**

Якість повітря в приміщенні впливає на самопочуття людини. У приміщеннях з комп'ютерною технікою, внаслідок того що третина споживаної нею енергії розсіюється у вигляді тепла, формуються специфічні умови мікроклімату а саме, підвищується температура до 26-27 °С, відносна вологість падає до 40%. Вміст двоокису вуглецю збільшується внаслідок знаходження великої кількості користувачів в приміщенні, водночас повітря деіонізується, а це збільшує кількість позитивних іонів і несприятливо впливає на працездатність, самопочуття, сприйняття нового матеріалу.

Деіонізація повітря пояснюється притягненням від'ємних іонів до екрану монітора, який має позитивний потенціал і відштовхуванням позитивних іонів (це стосується моніторів з електронно-вакуумними трубками). Медичними дослідженнями встановлено, що на життєдіяльність людини впливає не кількість іонів повітря, а співвідношення між позитивно і негативно зарядженими іонами. Згідно норм в приміщеннях з комп'ютерною технікою оптимальним вважається вміст 3000-5000 від'ємних іонів кисню в 1 см<sup>3</sup>

повітря, а позитивних 1500-3000 іонів. Концентрація від'ємних іонів біля монітору повинна бути не меншою 600 іон/ см<sup>3</sup>.

У виробничих приміщеннях на робочих місцях з ПК мають забезпечуватись оптимальні значення параметрів мікроклімату: температури, відносної вологості й швидкості руху повітря (ДСанПіН 3.3.2.007-98, ДСН 3.3.6.042-99) (табл. 5.1).

Таблиця 5.1. Норми мікроклімату для приміщень з комп'ютерною технікою

Пора року	Категорія робіт	Температура повітря, С°, не більше	Відносна вологість повітря. %	Швидкість руху повітря, м/с
Холодна	легка - 1 а	22-24	40-60	0.1
	легка - 1б	21 -23	40-60	0.1
Тепла	легка - 1 а	23-25	40-60	0.1
	легка - 1б	22-24	40-60	0,2

У приміщеннях з комп'ютерною технікою має бути забезпечений 3-кратний обмін повітря за годину. Для забезпечення постійних параметрів мікроклімату (температури, вологості, швидкості руху і чистоти повітря) в приміщеннях можуть бути встановлені побутові кондиціонери.

## 5.2. Вимоги до освітлення приміщень та робочих місць

Приміщення з комп'ютерною технікою повинні мати природне та штучне освітлення. При незадовільному освітленні зменшується продуктивність праці користувачів ПК [4]. Система освітлення повинна відповідати таким вимогам:

✓ освітленість на робочому місці повинна відповідати характеру зорової роботи, який визначається об'єктом розрізнення - найменшим розміром об'єкту, що розглядається на моніторі персонального комп'ютера;

- ✓ необхідно забезпечити достатньо рівномірне розподілення яскравості на робочій поверхні монітора, а також в межах навколишнього простору;
- ✓ на робочій поверхні повинні бути відсутні різкі тіні;
- ✓ в полі зору не повинно бути відблисків;
- ✓ величина освітленості повинна бути постійною під час роботи;
- ✓ оптимальної спрямованості світлового потоку і необхідного складу світла.

Природне освітлення в приміщеннях з комп'ютерною технікою повинно відповідати вимогам ДСанПіН 3.3.2.007-98 та ДБН В 2.5.28-2006. Природне освітлення має здійснюватись через світлові прорізи, орієнтовані переважно на північ або північний схід і забезпечувати коефіцієнт природної освітленості (КПО) не нижче 1,5 %. Для захисту від прямих сонячних променів, які створюють прямі та відбиті відблиски на поверхні екранів і клавіатури, повинні бути передбачені сонцезахисні пристрої, вікна повинні мати жалюзі або штори. Задовільне природне освітлення легше створити в невеликих приміщеннях на 5-8 робочих місць.

Штучне освітлення в приміщеннях з робочими місцями, обладнаними ПК, має здійснюватись системою загального рівномірного освітлення. Значення освітленості на поверхні робочого столу в зоні розміщення документів має становити 300-500 лк. Якщо ці значення освітленості неможливо забезпечити системою загального освітлення, допускається використовувати місцеве освітлення. При цьому світильники місцевого освітлення слід встановлювати таким чином, щоб не створювати відблисків на поверхні екрана, а освітленість екрана має не перевищувати 300 лк [1].

Як джерела світла в штучному освітленні мають застосовуватись переважно люмінісцентні лампи типу ЛБ. Допускається застосування ламп розжарювання у світильниках місцевого освітлення.

Система загального освітлення має становити суцільні або переривчасті лінії світильників, розташовані збоку від робочих місць (переважно ліворуч), паралельно лінії зору працюючих.

Для штучного освітлення нормованим параметром виступає  $E_{\min}$  – мінімальний рівень освітленості, та  $K_n$  – коефіцієнт пульсації світлового потоку, який не повинний бути більшим ніж 20%. У табл. 5.2 наведені норми штучного та природного освітлення виробничих приміщень.

Таблиця 5.2 – Норми штучного та природного освітлення виробничих приміщень

Характеристика зорових робіт	Найменший розмір об'єкта розпізнавання, мм	Розряд зорової роботи	Підрозряд зорової роботи	Штучне освітлення	Природне освітлення
				Освітленість, лк	КПО, %
				загальне освітлення	бокове освітлення
Середньої точності	0,5-1	IV	а	300	1.5
			б	200	
			в	200	
			г	200	
Малої точності	1-5	V	а	300	1,0
			б	200	
			в	200	
			г	200	
Груба	Більше 5	VI	–	200	1,0

Мінімальна освітленість встановлюється в залежності від розряду виконуваних зорових робіт. Для IV розряду зорових робіт вона складає 300...500 лк.

### 5.3. Аналіз стану шуму у приміщенні

Характеристика джерел шуму у приміщенні наведена в таблиці 5.3.

Таблиця 5.3 – Характеристика джерел шуму

№ п/п	Найменування	Джерело небезпеки	Причини небезпек	Наслідки небезпеки
1	Сервер Dell PowerEdge T330	Елемент охолодження сервера (куллер)	Шум елемента охолодження	Емоційна роздратованість працівника, що може призвести до зменшення продуктивності його роботи

Засоби та заходи захисту від впливу шуму наведено у таблиці 5.4.

Таблиця 5.4 – Засоби та заходи захисту від небезпеки

№ п/п	Вид заходу	Зміст заходу	Результат заходу
1	Технічні	підтримка серверної шафи у закритому вигляді; встановлення шумоізоляційної перегородки між серверною та робочою зоною; додаткова шумоізоляція серверної шафи за допомогою ущільнювачів дверей і швів та резинових підставок під ноги шафи.	Мінімізування рівня шуму
2	Організаційні	перерви для працівника; перевірка на належний стан куллерів та шумоізоляційних матеріалів перегородки і серверної шафи.	Запобігання впливу на організм працівника шуму
3	Експлуатаційні	своєчасний ремонт куллерів та заміна шумоізоляційних елементів	Захист від високого рівня шуму

#### **5.4. Інструкція з охорони праці під час створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур**

Нами запропоновано інструкцію з охорони праці під час створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур.

##### *1. Загальні положення*

1.1. Ця інструкція розроблена з метою забезпечення безпечних та здорових умов праці при створенні та використанні інформаційних технологій прогнозування урожайності сільськогосподарських культур (ІТПУС).

1.2. Дія інструкції поширюється на всіх працівників, які беруть участь у створенні та використанні ІТПУС, незалежно від їх посади, професії та кваліфікації.

1.3. Відповідальність за дотримання правил охорони праці, викладених у цій інструкції, несуть:

Роботодавець:

✓ Забезпечує безпечні та здорові умови праці, проводить інструктаж з охорони праці, організовує навчання та медичні огляди працівників.

Працівники:

✓ Дотримуються правил охорони праці, використовують засоби індивідуального захисту, проходять інструктаж та навчання з охорони праці.

##### *2. Вимоги до безпеки при роботі з комп'ютерним обладнанням*

2.1. Перед початком роботи з комп'ютером необхідно:

✓ Перевірити візуально справність комп'ютера та його периферійних пристроїв.

✓ Увімкнути комп'ютер у мережу живлення.

✓ Завантажити операційну систему.

2.2. Під час роботи з комп'ютером необхідно:



- ✓ Сидіти прямо, тримаючи спину рівно.
- ✓ Розташувати екран комп'ютера на відстані 50-70 см від очей.
- ✓ Робити перерви в роботі за комп'ютером через кожну годину роботи на 10-15 хвилин.

- ✓ Не допускати перевантажень очей.
- ✓ Не їсти та не пити за робочим столом.
- ✓ Не класти на комп'ютер сторонні предмети.
- ✓ Не допускати попадання рідини на комп'ютер.
- ✓ Не торкатися до оголених проводів та інших елементів комп'ютера, які знаходяться під напругою.

2.3. Після закінчення роботи з комп'ютером необхідно:

- ✓ Вимкнути комп'ютер з мережі живлення.
- ✓ Прибрати робоче місце.

### *3. Вимоги до безпеки при роботі з програмним забезпеченням*

3.1. Перед початком роботи з програмним забезпеченням необхідно:

- ✓ Ознайомитися з інструкцією з експлуатації програмного забезпечення.

- ✓ Переконаватися, що програмне забезпечення ліцензійне.

3.2. Під час роботи з програмним забезпеченням необхідно:

- ✓ Не використовувати піратське програмне забезпечення.
- ✓ Не встановлювати на комп'ютер програмне забезпечення з неперевіраних джерел.

- ✓ Регулярно оновлювати програмне забезпечення.

- ✓ Використовувати антивірусне програмне забезпечення.

- ✓ Не розголошувати конфіденційну інформацію.

3.3. Після закінчення роботи з програмним забезпеченням необхідно:

- ✓ Закрити всі програми.
- ✓ Вимкнути комп'ютер з мережі живлення.

### *4. Вимоги до безпеки при роботі на відкритому повітрі*

4.1. Перед початком роботи на відкритому повітрі необхідно:

- ✓ Ознайомитися з прогнозом погоди.
- ✓ Одягнутися відповідно до погодних умов.
- ✓ Взяти з собою необхідні засоби індивідуального захисту.

4.2. Під час роботи на відкритому повітрі необхідно:

- ✓ Дотримуватися правил безпеки при роботі з електроінструментом.
- ✓ Бути обережними при роботі з хімічними речовинами.
- ✓ Захищатися від сонячних променів.
- ✓ Пити багато води.
- ✓ Не працювати в екстремальних погодних умовах.

4.3. Після закінчення роботи на відкритому повітрі необхідно:

- ✓ Зняти спецодяг та засоби індивідуального захисту.
- ✓ Прийняти душ.
- ✓ Переодягнутися в чистий одяг.

### *5. Навчання та інструктаж з охорони праці*

5.1. Всі працівники, які беруть участь у створенні та використанні ІТПУС, повинні пройти:

- ✓ Вступний інструктаж з охорони праці;
- ✓ Первинний інструктаж з охорони праці на робочому місці;
- ✓ Інструктаж з охорони праці перед початком роботи;
- ✓ Цільовий інструктаж з охорони праці при виконанні робіт з підвищеною небезпекою;

- ✓ Повторний інструктаж з охорони праці;
- ✓ Позаплановий інструктаж з охорони праці.

5.2. Інструктаж з охорони праці проводять:

- ✓ Інженер з охорони праці;
- ✓ Особа, відповідальна за охорону праці на робочому місці.

Дотримання правил охорони праці, викладених у цій інструкції, є запорукою безпечного та здорового робочого середовища, а також запобігання нещасних випадків та професійних захворювань.

## ВИСНОВКИ І ПРОПОЗИЦІЇ

В умовах сучасного розвитку технологій, інформаційні технології відіграють все більш важливу роль у сільському господарстві. Метою даної кваліфікаційної роботи є розробка та впровадження інформаційної технології для прогнозування врожайності сільськогосподарських культур. У процесі виконання роботи розглянуто сучасні методи та підходи до прогнозування врожайності, зокрема із застосуванням методів машинного навчання, аналізу великих даних та проектування інтелектуальних інформаційних технологій.

Прогнозування врожайності сільськогосподарських культур є складним та багатограним процесом, що вимагає врахування широкого спектра факторів (рис. 1.1). Нами розглянуто особливості впливу природних факторів на процес прогнозування урожайності сільськогосподарських культур. Прогнозування врожайності значною мірою залежить від природних умов, таких як клімат, погодні умови, тип ґрунту та наявність водних ресурсів. Зміни клімату можуть суттєво впливати на вегетаційний період та стан рослин, що вимагає точних і своєчасних даних для коректного прогнозування.

Пропонується використання на основі Sentinel-2 нормалізованого індексу різниці рослинності (NDVI) як проксі для оцінки врожайності шляхом співвіднесення його з ССЕ піксельного рівня є звичайною практикою в сільському господарстві. NDVI є надійним показником урожайності, оскільки він відображає здоров'я та ріст рослинності. Також відома методологія, прийнята для картографування типів культур. Розроблено підхід до оцінки врожайності за допомогою DSSAT та її інтеграції з дистанційним зондуванням.

Заслуговує на увагу використання алгоритмів машинного навчання для прогнозування урожайності сільськогосподарських культур. Машинне навчання набуло популярності в сільському господарстві завдяки своєму успіху в інших галузях, таких як медицина, біоінформатика та обробка природної мови. У роботах розглядають застосування машинного навчання в сільському господарстві. Запропоновано підхід до моделювання врожайності який

використовує напівпараметричний варіант глибокої нейронної мережі, що може одночасно враховувати складні нелінійні зв'язки у високорозмірних наборах даних, а також відому параметричну структуру та неспостережувану поперечну неоднорідність.

Відома піраміда знань, що пов'язує дані з інформацією, знаннями і мудрістю, в якій дані є сировиною для розробки додатків, спрямованих на прийняття рішень. Вона лежить в основі використання технологій машинного навчання для вирішення різних завдань.

Описано взаємозв'язки між складовими інформаційних технологій, що відображають потік даних та інформації через рівні моделювання, синтезу та інтерфейсу до кінцевих користувачів з роллю різних акторів. Оскільки сільськогосподарські підприємства відрізняються за розміром, наявністю робочої сили, ґрунтами та іншими характеристиками, запропоновано використовувати інструменти NextGen, щоб допомогти адаптувати поради для кожного господарства.

Встановлено, що машинне навчання (ML) та штучний інтелект (AI) стали невід'ємними компонентами сучасних систем прогнозування врожайності. Отже, розробка інформаційної технології для точного прогнозування врожайності сільськогосподарських культур має велике значення для розвитку сучасного сільського господарства. Використання сучасних інформаційних технологій дозволить підвищити точність прогнозування врожайності, оптимізувати процеси та забезпечити стале зростання аграрного сектора.

Нами розроблена узагальнена структурна схема архітектури інформаційної технології прогнозування урожайності сільськогосподарських культур, яка передбачає 7 блоків, базується на моделі машинного навчання (навчених моделей RandomForestRegressor), представлена на рис. 2.1.

Нами виконано детальний аналіз вимог до інформаційної технології прогнозування урожайності сільськогосподарських культур. Аналіз включає в себе розгляд потреб користувачів, визначення функціональних та

нефункціональних вимог до системи, а також формулювання основних функцій та характеристик системи (рис. 2.2).

Нами проведений аналіз та вибір засобів та інструментів, необхідних для реалізації інформаційної технології прогнозування урожайності (рис. 2.3). Це включатиме вибір програмних мов, баз даних, бібліотек машинного навчання та інших інструментів. Пропонується використовувати мову програмування – Python. Основною перевагою цієї мови є простота у використанні, велика кількість бібліотек для аналізу даних та машинного навчання (numpy, pandas, scikit-learn). Вона вміщує бібліотеки NumPy, Pandas, Scikit-learn та Matplotlib для візуалізації. Причому SQL використовується для управління та взаємодії з реляційними базами даних. Щодо вибору баз даних, то пропонується PostgreSQL. Вибрана реляційна база даних забезпечує зберігання зібраних даних про погоду, урожайність, використання пестицидів тощо. Нами вибрано бібліотеку машинного навчання – Scikit-learn. Це основна бібліотека для реалізації алгоритмів машинного навчання, включаючи RandomForestRegressor.

Нами виконано проектування інтерфейсів користувача та інтеграційних інтерфейсів для взаємодії з системою прогнозування урожайності. Це включатиме розробку UI/UX дизайну, а також API для зовнішніх систем. Запропонована структура інтерфейсу та написаний код дозволяють створити функціональний та зручний веб-додаток для прогнозування урожайності сільськогосподарських культур. Цей інтерфейс забезпечує користувачам можливість легко вводити необхідні дані, переглядати результати прогнозування та проводити аналітичний аналіз.

Нами написано код для реалізації API на Flask (Python). Цей код створює веб-додаток за допомогою фреймворку Flask для прогнозування урожайності сільськогосподарських культур на основі моделі машинного навчання, завантаженої з файлу. Це забезпечує простий інтерфейс для взаємодії з моделлю машинного навчання для прогнозування урожайності. Він включає маршрути для отримання прогнозів, додавання даних та проведення аналітики.

Нами розглядається набір даних, що використовується для прогнозування врожайності сільськогосподарських культур за допомогою алгоритмів машинного навчання (рис. 3.1). Набір даних включає різноманітні кліматичні змінні та інші чинники, що впливають на врожайність. У нашому наборі даних наявно 7 атрибутів.

Нами виконано перевірку набору даних на наявність пропущених значень (рис. 3.2). Отриманий набір даних немає пропущених значень. У ньому є 12254 екземпляри даних, які характеризують є 3 унікальні значення – «Potatoes» (Картопля), «Maize» (Кукурудза) та «Wheat» (Пшениця).

Нами написано код, який дав можливість створити графіки впливу кількості опадів (рис. 3.3) та температури (рис. 3.5) на урожайність сільськогосподарських культур. Встановлено, що урожайності сільськогосподарських культур залежить від кількості опадів по різному для кожної із досліджуваних сільськогосподарських культур. При цьому за опадів більше 2000 мм / рік значно знижується врожайність. Найбільша врожайність спостерігається за 600...1800 мм/рік. найвища урожайність картоплі припадає на середньорічну температуру 7...15 0C, водночас пшениці – 10...15 0C. За середньорічної температури 22 0C урожайність усіх культур значно знижується.

Насамперед нами здійснено імпорт потрібних модулів та вибору моделей (рис. 4.3). Після цього виконуємо поділ даних на тренувальні та тестові набори (рис. 4.4). Наступним кроком є навчання моделей прогнозування урожайності сільськогосподарських культур (рис. 4.5).

Після навчання моделей виконується прогнозування значень на тестовому наборі `x_test` за допомогою методу `predict()`. Встановлено, що найкращою моделлю з цих результатів є `RandomForestRegressor`, оскільки вона має найвище значення скоригованого R-квадрату та найнижче значення середньоквадратичного відхилення (RMSE). `RandomForestRegressor` має трохи кращі показники за всіма цими показниками, що можна вважати її найкращою моделлю для прогнозування врожайності сільськогосподарських культур. Для

RandomForestRegressor числові значення метрики ефективності такі. Скоригований R-квадрат становить 0,981, R-Squared (Коефіцієнт детермінації) – 0,982, RMSE (Середньоквадратичне відхилення) – 13.142 та тривалість навчання моделі – 5.939 секунд.

Оцінення процесу навчання моделі RandomForestRegressor виконаємо на основі побудованого графіка, який відображає зміну точності моделі на збільшення розміру процесу навчального набору даних (рис. 4.6). За допомогою графіка (рис. 4.7) бачимо, модель RandomForestRegressor добре прогнозує значення урожайності сільськогосподарських культур. Точки на графіку розсіювання розташовані близько до діагональної лінії, це означає, що модель RandomForestRegressor успішно прогнозує значення урожайності на основі вхідних даних.

У роботі розроблено заходи із безпеки під час створення та використання інформаційних технологій прогнозування урожайності сільськогосподарських культур.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Гандзюк М. П., Желібо Е. П., Халімовський М. О. Основи охорони праці / За ред. Гандзюка М. П. - К.: Каравела 2003. 405 с.
2. ДСТУ 2293-99 Охорона праці. Терміни та визначення основних понять. Київ, 1999.
3. Жидецький В.Ц. Основи охорони праці: Підручник. К.: Основа, 2002. 320 с.
4. Москальова В. М. Основи охорони праці. Підручник. Київ: ВД Професіонал, 2005. 666 с.
5. Програмування числових методів мовою Python : підруч. / А. В. Анісімов, А. Ю. Дорошенко, С. Д. Погорілий, Я. Ю. Дорогий ; за ред. А. В. Анісімова. К. : Видавничо-поліграфічний центр «Київський університет», 2014. 640 с.
6. Супутникові знімки, індекс NDVI. URL: <https://www.soft.farm/uk/crop-husbandry/ndvi>
7. Tryhuba A., Boyarchuk V., Tryhuba I., Ftoma O., Padyuka R., Rudynets M., Forecasting the Risk of the Resource Demand for Dairy Farms Basing on Machine Learning, Proceedings of the 2nd International Workshop on Modern Machine Learning Technologies and Data Science (MoMLLeT+DS 2020). Volume I: Main Conference. Lviv-Shatsk, Ukraine, June 2-3, 2020. pp.327-340.
8. Ткачук К. Н., Халімовський М. О., Зацарний В.В., та інші. Основи охорони праці: Підручник. К.: Основа, 2006. 444 с.
9. Ahamed A.T.M.S., Mahmood N.T., Hossain N., Kabir M.T., Das K., Rahman F., Rahman R.M. Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh, 2015. IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPDP 2015 - Proceedings (2015), 10.1109/SNPDP.2015.7176185



10. Attri, S.; Tyagi, A. Climate Profile of India; Environment Monitoring and Research Center, India Meteorology Department: New Delhi, India, 2010.
11. Crane-Droesch A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environ. Res. Lett.*, 13 (2018), p. 114003.
12. GISGeography. URL: <https://gisgeography.com/ndvi-normalized-difference-vegetation-index/>
13. Gumma, M.K.; Thenkabail, P.S.; Panjala, P.; Teluguntla, P.; Yamano, T.; Mohammed, I. Multiple agricultural cropland products of south asia developed using landsat-8 30 m and modis 250 m data using machine learning on the google earth engine (gee) cloud and spectral matching techniques (SMTs) in support of food and water security. *GISci. Remote Sens.* 2022, 59, 1048–1077.
14. Holzworth D.P. et al. APSIM–evolution towards a new generation of agricultural systems simulation. *Environ. Model Softw.*, 62 (2014), pp. 327-350.
15. Janssen, S., Linking models for assessing agricultural land use change. *Comput. Electron. Agric.* 2011, 76, 148–160.
16. Jha, M.N.; Kumar, A.; Dubey, S.; Pandey, A. Yield estimation of rice crop using semi-physical approach and remotely sensed data. In *Geospatial Technologies for Resources Planning and Management*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 331–349.
17. Koval N., Tryhuba A., Kondysiuk I., Tryhuba I., Boiarchuk O. Forecasting the Fund of Time for Performance of Works in Hybrid Projects Using Machine Training Technologies. *MoMLLeT+ DS*, 196-206.
18. Liakos K., Busato P., Moshou D., Pearson S., Bochtis D. Machine learning in agriculture: a review *Sensors*, 18, 2018, P. 2674.
19. Lokers, R., Knapen, R., Janssen, S., van Randen, Y., Jansen, J., 2016. Analysis of big data technologies for use in agro-environmental science. *Environ. Model. Softw.* 84, 494–504.
20. Malanchuk O., Tryhuba A., Tryhuba I., Sholudko R., Pankiv O., A Neural Network Model-based Decision Support System for Time Management in Pediatric

Diabetes Care Projects. International Scientific and Technical Conference on Computer Sciences and Information Technologies, 2023.

21. Malanchuk O.M., Triguba A.M., Pankiv O.V., Sholudko R.Y. Architecture of an intelligent information system for forecasting components of medical projects. *Applied aspects of information technology* 6 (4), 376-390-376-390.

22. Matsumura K., Gaitan C.F., Sugimoto K., Cannon A.J., Hsieh W.W. Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. *J. Agric. Sci.*, 153 (3) (2015), pp. 399-410.

23. Meraj, G.; Kanga, S.; Ambadkar, A.; Kumar, P.; Singh, S.K.; Farooq, M.; Johnson, B.A.; Rai, A.; Sahu, N. Assessing the yield of wheat using satellite remote sensing-based machine learning algorithms and simulation modeling. *Remote Sens.* 2022, 14, 3005.

24. Paudel, D., Boogaard, H., de Wit, A., Pylianidis, C., Athanasiadis, I.N. Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, 2021, 187, 103016.

25. Pazhanivelan, S.; Geethalakshmi, V.; Tamilmounika, R.; Sudarmanian, N.; Kaliaperumal, R.; Ramalingam, K.; Sivamurugan, A.; Mrunalini, K.; Yadav, M.K.; Quicho, E.D. Spatial rice yield estimation using multiple linear regression analysis, semi-physical approach and assimilating sar satellite derived products with DSSAT crop simulation model. *Agronomy* 2022, 12, 2008.

26. Tripathy, R.; Chaudhary, K.; Nigam, R.; Manjunath, K.; Chauhan, P.; Ray, S.; Parihar, J. Operational semi-physical spectral-spatial wheat yield model development. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 2014, 40, 977–982.

27. Tryhuba A., Boyarchuk V., Tryhuba I., Francik S., Rudynets M., Method and software of planning of the substantial risks in the projects of production of raw material for biofuel, *CEUR Workshop Proceedings*, 2020, 2565, pp. 116–129.

28. Tryhuba A., Kondysiuk I., Tryhuba I., Boiarchuk O., Tatomyr A., Intellectual information system for formation of portfolio projects of motor transport enterprises, *CEUR Workshop Proceedings*, 2022, 3109, pp. 44–52.

29. Tryhuba A., Kotenko V. Intelligent information system for resource planning in grain crops delivery projects on the basis of machine learning. IEEE 18th International Conference on Computer Science and Information Technologies (CSIT) 2023. P. 1-4.

30. Tryhuba A., Koval N., Tryhuba I., Boiarchuk O., Application of Sarima Models in Information Systems Forecasting Seasonal Volumes of Food Raw Materials of Procurement on the Territory of Communities. CEUR Workshop Proceedings, 2022, 3295, pp. 64–75.

31. Tryhuba A., Padyuka R., Tymochko V., Lub P., Mathematical model for forecasting product losses in crop production projects. CEUR Workshop Proceedings, 2022, 3109, pp. 25–31.

32. Tryhuba A., Tryhuba I., Ftoma O., Boyarchuk O., Method of quantitative evaluation of the risk of benefits for investors of fodder-producing cooperatives. International Scientific and Technical Conference on Computer Sciences and Information Technologies, 2019, 3, pp. 55–58, 8929788.

## Додатки

## Додаток А

### Фрагмент синтаксису інтерфейсу користувача

#### Головна панель (HTML)

```

<!DOCTYPE html>
<html lang=«en»>
<head>
  <meta charset=«UTF-8»>
  <title>Прогнозування урожайності</title>
  <link rel=«stylesheet» href=«styles.css»>
</head>
<body>
  <div class=«navbar»>
    <a href=«#home»>Головна</a>
    <a href=«#data-entry»>Введення даних</a>
    <a href=«#prediction-results»>Результати прогнозування</a>
    <a href=«#analytics»>Аналітика</a>
    <a href=«#reports»>Звіти</a>
  </div>
</body>
</html>

<div id=«data-entry»>
  <h2>Введення даних для прогнозування</h2>
  <form id=«dataForm»>
    <label for=«region»>Область:</label>
    <input type=«text» id=«region» name=«region» required><br>

    <label for=«year»>Рік:</label>
    <input type=«number» id=«year» name=«year» required><br>

    <label for=«rainfall»>Середня кількість опадів (мм):</label>
    <input type=«number» id=«rainfall» name=«rainfall» required><br>

    <label for=«pesticides»>Тонни пестицидів:</label>
    <input type=«number» id=«pesticides» name=«pesticides» required><br>

    <label for=«temperature»>Середня температура (°C):</label>
    <input type=«number» id=«temperature» name=«temperature» required><br>

    <button type=«submit»>Прогнозувати</button>
  </form>
</div>

<div id=«prediction-results»>
  <h2>Результати прогнозування</h2>
  <table>
    <thead>
      <tr>
        <th>Область</th>
        <th>Рік</th>
        <th>Урожайність (ц/га)</th>
      </tr>
    </thead>
    <tbody id=«resultsTable»>
      <!-- Динамічне додавання рядків таблиці -->
    </tbody>
  </table>
  <canvas id=«resultsChart»></canvas>
</div>

```

```

<div id=«analytics»>
  <h2>Аналітика</h2>
  <div id=«charts»>
    <canvas id=«analysisChart1»></canvas>
    <canvas id=«analysisChart2»></canvas>
  </div>
</div>

<div id=«reports»>
  <h2>Звіти</h2>
  <button id=«generatePdfReport»>Завантажити PDF звіт</button>
  <button id=«generateExcelReport»>Завантажити Excel звіт</button>
</div>

```

### Форма введення даних (HTML)

```

<div id=«data-entry»>
  <h2>Введення даних для прогнозування</h2>
  <form id=«dataForm»>
    <label for=«region»>Область:</label>
    <input type=«text» id=«region» name=«region» required><br>

    <label for=«year»>Рік:</label>
    <input type=«number» id=«year» name=«year» required><br>

    <label for=«rainfall»>Середня кількість опадів (мм):</label>
    <input type=«number» id=«rainfall» name=«rainfall» required><br>

    <label for=«pesticides»>Тонни пестицидів:</label>
    <input type=«number» id=«pesticides» name=«pesticides» required><br>

    <label for=«temperature»>Середня температура (°C):</label>
    <input type=«number» id=«temperature» name=«temperature» required><br>

    <button type=«submit»>Прогнозувати</button>
  </form>
</div>

```

### Результати прогнозування (HTML)

```

<div id=«prediction-results»>
  <h2>Результати прогнозування</h2>
  <table>
    <thead>
      <tr>
        <th>Область</th>
        <th>Рік</th>
        <th>Урожайність (hg/ra)</th>
      </tr>
    </thead>
    <tbody id=«resultsTable»>
      <!-- Динамічне додавання рядків таблиці -->
    </tbody>
  </table>
  <canvas id=«resultsChart»></canvas>
</div>

```

### Аналітика (HTML)

```

<div id=«analytics»>
  <h2>Аналітика</h2>
  <div id=«charts»>

```

```

        <canvas id=«analysisChart1»></canvas>
        <canvas id=«analysisChart2»></canvas>
    </div>
</div>

```

## Звіти (HTML)

```

<div id=«reports»>
    <h2>Звіти</h2>
    <button id=«generatePdfReport»>Завантажити PDF звіт</button>
    <button id=«generateExcelReport»>Завантажити Excel звіт</button>
</div>

```

## Код реалізації API на Flask (Python)

```

from flask import Flask, request, jsonify
import joblib
import pandas as pd

app = Flask(__name__)

# Завантаження навченої моделі з абсолютним шляхом
model = joblib.load(...\random_forest_model.pkl')

@app.route('/api/v1/prediction', methods=['POST'])
def predict():
    data = request.json
    input_data = pd.DataFrame([data])
    prediction = model.predict(input_data)
    return jsonify({'prediction': prediction[0]})

@app.route('/api/v1/data', methods=['POST'])
def add_data():
    data = request.json
    # Логіка збереження даних у базу даних
    return jsonify({'status': 'success'})

@app.route('/api/v1/analysis', methods=['GET'])
def analysis():
    # Логіка отримання аналітичних даних
    analysis_data = {'some_key': 'some_value'}
    return jsonify(analysis_data)

if __name__ == '__main__':
    app.run(debug=True)

```