

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

# **КВАЛІФІКАЦІЙНА РОБОТА**

другого (магістерського) рівня вищої освіти

на тему: **«Прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання»**

Виконав: студент групи Іт-61

Спеціальності 126 «Інформаційні системи та технології»

(шифр і назва)

Жидачевський Дмитро Мар'янович

(Прізвище та ініціали)

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

(Прізвище та ініціали)

Рецензент: к.т.н., доцент Шарибура А.О.

(Прізвище та ініціали)

**ДУБЛЯНИ-2024**

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

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

«ЗАТВЕРДЖУЮ»

Завідувач кафедри \_\_\_\_\_

д.т.н., проф. А.М. Тригуба

« \_\_\_\_ » \_\_\_\_\_ 2024 р.

### ЗАВДАННЯ

на кваліфікаційну роботу студенту

Жидачевському Дмитру Мар'яновичу

1. Тема роботи: «Прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання»

Керівник роботи Тригуба Анатолій Миколайович, професор  
затверджені наказом по університету від 12.09.2024 року № 616/к-с.

2. Строк подання студентом роботи 06.12.2024 р.

3. Вихідні дані до роботи: дані для машинного навчання; методика алгоритми машинного навчання для прогнозування.

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

Вступ.

1. Аналіз стану питання в теорії та практиці.

2. Вибір моделей та підготовка даних для прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання.

3. Результати розробки моделей прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання.

4. Охорона праці та безпека у надзвичайних ситуаціях.

5. Визначення ефективності від запропонованих розробок.

Висновки та пропозиції.

Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових слайдів): аналіз стану питання в теорії та практиці; вибір моделей та підготовка даних для прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання; результати розробки моделей прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання; економічна ефективність.

6. Консультанти з розділів:

| Розділ     | Прізвище, ініціали та посада консультанта  | Підпис, дата   |                  |
|------------|--|----------------|------------------|
|            |  | завдання видав | завдання прийняв |
| 1, 2, 3, 5 | <i>Тригуба А.М., д.т.н., професор кафедри інформаційних технологій</i>                     |                |                  |
| 4          | <i>Городецький І.М., доцент кафедри фізики, інженерної механіки та безпеки виробництва</i> |                |                  |

7. Дата видачі завдання

12 вересня 2024 р.

#### Календарний план

| № з/п | Назва етапів кваліфікаційної роботи   | Терміни виконання етапів роботи | Примітка |
|-------|---|---------------------------------|----------|
| 1     | <i>Написання першого розділу</i>  | <i>12.09-20.09.24</i>           |          |
| 2     | <i>Виконання другого розділу та аркушів ілюстраційного матеріалу до нього</i>                       | <i>21.09-14.10.24</i>           |          |
| 3.    | <i>Виконання третього розділу та аркушів ілюстраційного матеріалу до нього</i>                      | <i>15.10-10.11.24</i>           |          |
| 4.    | <i>Написання розділу «Охорона праці та безпека у надзвичайних ситуаціях»</i>                        | <i>11.11-20.11.24</i>           |          |
| 5.    | <i>Оцінення ефективності запропонованої системи</i>   | <i>21.11-30.30.24</i>           |          |
| 6.    | <i>Завершення оформлення розрахунково-пояснювальної записки та аркушів ілюстраційного матеріалу</i> | <i>01-04.12.24</i>              |          |

Студент \_\_\_\_\_ Жидачевський Д.М.  
(підпис)

Керівник роботи \_\_\_\_\_ Тригуба А.М.  
(підпис)

УДК 004.8:631.1

Прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання.

Жидачевський Д.М. Кафедра інформаційних технологій – Дубляни, ЛНУП, 2024.

Кваліфікаційна робота: 68 с. текст. част., 14 рис., 4 табл., 12 арк. ілюстраційного матеріалу, 42 джерела.

Виконано аналіз стану предметної області та прогнозування врожайності сільськогосподарських культур. Обґрунтована доцільність використання машинного навчання для прогнозування врожайності сільськогосподарських культур. Здійснено аналіз використання алгоритмів ml для прогнозування врожайності сільськогосподарських культур. Обґрунтована доцільність прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання.

Подана схема використовуваних моделей для дослідження прогнозування врожайності сільськогосподарських культур. Наведено аналіз використовуваних моделей прогнозування врожайності сільськогосподарських культур (множинна лінійна регресія, екстремальне градієнтне підсилювання XGboost, регресія випадкового лісу RFR, модель ансамблю, градієнтне підсилювання Gradient Boosting. Подано метрики оцінювання. Здійснено підготовку даних для прогнозування врожайності сільськогосподарських культур. Створено моделі прогнозування врожайності сільськогосподарських культур. Подано результати використання створених моделей для прогнозування врожайності сільськогосподарських культур. Наведено результати визначення ефективної моделі для прогнозування врожайності сільськогосподарських культур.

Ключові слова: машинне навчання, прогнозування, врожайність, сільськогосподарські культури.

## ЗМІСТ

|  |    |
|--|----|
| ВСТУП .....  | 7  |
| РОЗДІЛ 1. АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ.....   | 10 |
| 1.1. Аналіз стану предметної області та прогнозування врожайності сільськогосподарських культур .....  | 10 |
| 1.2. Доцільність використання машинного навчання для прогнозування врожайності сільськогосподарських культур.....                                      | 13 |
| 1.3. Аналіз використання алгоритмів ml для прогнозування врожайності сільськогосподарських культур .....   | 16 |
| 1.4. Обґрунтування доцільності прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання.....                | 21 |
| РОЗДІЛ 2. ВИБІР МОДЕЛЕЙ ТА ПІДГОТОВКА ДАНИХ ДЛЯ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР ІЗ ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ .. | 24 |
| 2.1. Схема використовуваних моделей для дослідження прогнозування врожайності сільськогосподарських культур.....                                       | 24 |
| 2.2. Аналіз використовуваних моделей прогнозування врожайності сільськогосподарських культур .....   | 25 |
| 2.2.1. Множинна лінійна регресія.....  | 25 |
| 2.2.2. Екстремальне градієнтне підсилювання XGboost .....  | 25 |
| 2.2.3. Регресія випадкового лісу RFR.....  | 26 |
| 2.2.4. Модель ансамблю.....  | 28 |
| 2.2.5. Градієнтне підсилювання Gradient Boosting .....   | 29 |
| 2.3. Метрики оцінювання .....  | 31 |
| 2.5. Підготовка даних для прогнозування врожайності сільськогосподарських культур  | 32 |

|   |    |
|---|----|
| РОЗДІЛ 3. РЕЗУЛЬТАТИ РОЗРОБКИ МОДЕЛЕЙ ПРОГНОЗУВАННЯ<br>ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР ІЗ<br>ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ.....  | 39 |
| 3.1. Створення моделей прогнозування врожайності сільськогосподарських<br>культур .....   | 39 |
| 3.2. Результати використання створених моделей для прогнозування<br>врожайності сільськогосподарських культур.....  | 42 |
| 3.3. Результати визначення ефективної моделі для прогнозування врожайності<br>сільськогосподарських культур .....   | 46 |
| РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ<br>СИТУАЦІЯХ .....  | 49 |
| 4.1. Небезпечні і шкідливі виробничі чинники під час створення моделей<br>машинного навчання.....   | 49 |
| 4.2. Заходи щодо попередження небезпечних і шкідливих<br>виробничих чинників .....  | 51 |
| 4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час<br>монтажу інтелектуальної інформаційної системи прогнозування врожайності<br>сільськогосподарських культур із використанням методів<br>машинного навчання..... | 52 |
| 4.4. Розробка заходів із забезпечення безпеки під час<br>надзвичайних ситуацій .....  | 55 |
| РОЗДІЛ 5. ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ<br>МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ<br>ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР.....   | 57 |
| ВИСНОВКИ І ПРОПОЗИЦІЇ.....  | 59 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....   | 64 |

## ВСТУП

Сучасні умови розвитку аграрного сектору ставлять перед виробниками складні завдання, пов'язані з підвищенням ефективності та стабільності виробництва сільськогосподарської продукції [32]. Одним із ключових факторів, що впливають на досягнення цих цілей, є точне прогнозування врожайності сільськогосподарських культур. Традиційні методи прогнозування часто виявляються недостатньо точними через складність врахування численних змінних факторів, таких як погодні умови, типи ґрунтів, технології вирощування та інші агротехнічні параметри. Використання сучасних підходів на основі машинного навчання дозволяє суттєво підвищити точність прогнозів завдяки аналізу великих обсягів даних і виявленню прихованих закономірностей.

На сьогоднішній день, завдяки розвитку інформаційних технологій, аграрна галузь має можливість впроваджувати інноваційні підходи до обробки даних. Зокрема, машинне навчання (ML) стало важливим інструментом у моделюванні та прогнозуванні, що забезпечує ефективне використання наявних ресурсів і підвищення продуктивності [29, 30]. Методи машинного навчання дозволяють аналізувати історичні дані, виявляти складні взаємозв'язки між змінними та адаптувати моделі до різних умов виробництва. Це є особливо важливим для сільського господарства, де високий рівень невизначеності та численні ризики можуть значно вплинути на обсяги врожаю.

Метою даної роботи є розробка підходів до прогнозування врожайності сільськогосподарських культур з використанням методів машинного навчання, що забезпечують підвищення точності прогнозів і сприяють оптимізації процесів планування та управління виробництвом.

Об'єктом дослідження є процеси формування врожайності основних сільськогосподарських культур та їх прогнозування. Предметом дослідження є методи машинного навчання, що використовуються для аналізу та

прогнозування врожайності сільськогосподарських культур, а також вплив різних факторів на результати прогнозування.

Завдання дослідження включають:

1. Огляд сучасних підходів до прогнозування врожайності сільськогосподарських культур;
2. Аналіз і вибір найбільш ефективних методів машинного навчання для цієї задачі;
3. Розробка моделі прогнозування врожайності на основі обраних методів;
4. Проведення експериментальних досліджень та оцінка ефективності запропонованої моделі;
5. Формулювання рекомендацій щодо впровадження розробленої моделі у практичну діяльність аграрних підприємств.

У процесі дослідження використовувалися такі методи – аналіз і синтез наукової літератури, методи машинного навчання (регресійні моделі, нейронні мережі, метод опорних векторів), математичне моделювання та статистичні методи для оцінки точності прогнозів.

Наукова новизна роботи полягає в адаптації сучасних методів машинного навчання для специфічних умов прогнозування врожайності сільськогосподарських культур, що дозволяє забезпечити високу точність і надійність прогнозів. Використання алгоритмів машинного навчання, таких як нейронні мережі, метод опорних векторів, регресійні моделі та ансамблеві методи, сприяє розширенню можливостей аграрного виробництва у плануванні та прийнятті рішень.

Практична значущість полягає в тому, що результати даного дослідження можуть бути використані аграрними підприємствами для покращення процесів управління, планування сівозмін, оцінки ефективності використання земельних ресурсів та зменшення впливу несприятливих факторів. Прогнозування врожайності на основі машинного навчання дозволяє зменшити ризики, пов'язані з непередбачуваними змінами умов і забезпечити стабільність



виробництва, що в кінцевому підсумку сприяє підвищенню рентабельності аграрного бізнесу.

Таким чином, розробка ефективних методів прогнозування врожайності із використанням машинного навчання є важливим кроком до впровадження технологій цифрової трансформації у сільському господарстві та підвищення його конкурентоспроможності у сучасному світі.

## РОЗДІЛ 1.

### АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ

#### 1.1. Аналіз стану предметної області та прогнозування врожайності сільськогосподарських культур

Аграрний сектор є важливою складовою економіки багатьох країн світу, включаючи Україну. Сільське господарство забезпечує населення продовольством та є джерелом значної частини національного доходу. У зв'язку з глобальними змінами клімату, збільшенням населення та іншими викликами, перед сільським господарством стоїть завдання підвищення ефективності виробництва та стабільності врожайності. Одним із ключових елементів управління аграрним виробництвом є точне прогнозування врожайності, яке дає змогу оптимізувати ресурси, знижувати ризики та забезпечувати стабільність постачання продукції.

Україна традиційно є однією з провідних аграрних країн, відомою своїми значними обсягами виробництва зернових та олійних культур. Зокрема, українські чорноземи є одними з найродючіших у світі, що створює сприятливі умови для сільськогосподарського виробництва. Проте, врожайність у країні піддається суттєвому впливу зовнішніх чинників, таких як погодні умови, захворювання рослин, технологічний рівень агротехніки та зміни клімату (рис. 1.1).

Традиційні методи прогнозування, засновані на аналізі історичних даних та експертних оцінках, не завжди забезпечують необхідну точність. Це створює потребу у впровадженні сучасних методів, таких як машинне навчання, які дозволяють враховувати велику кількість змінних факторів та підвищувати точність прогнозів.

В Україні вже спостерігаються перші успіхи у застосуванні інформаційних технологій для прогнозування врожайності. Дослідження показують, що використання методів машинного навчання дозволяє аграріям

краще адаптуватися до змінних умов та швидше реагувати на можливі загрози [10].

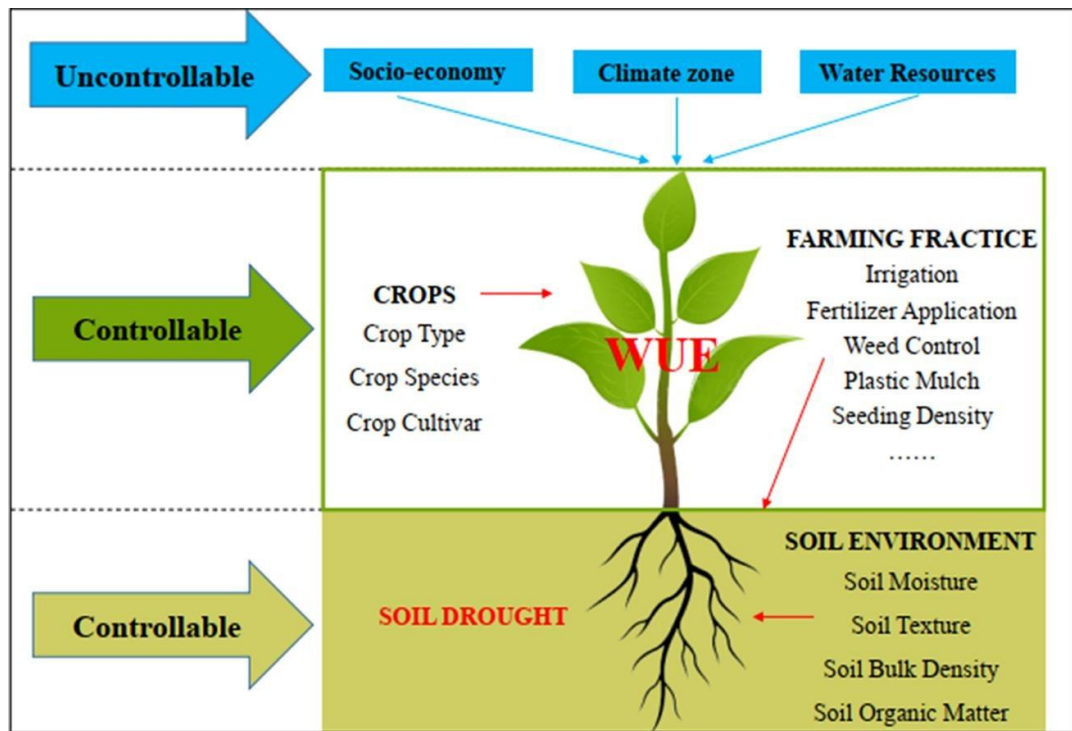


Рисунок 1.1 – Вплив чинників на врожайність сільськогосподарських культур [10]

Проте, масштабне впровадження таких технологій ще залишається обмеженим через недостатню кількість даних, ресурси для обробки великих обсягів інформації та брак фахівців у сфері аналітики даних.

На світовій арені прогнозування врожайності сільськогосподарських культур є однією з ключових задач для багатьох країн. У Сполучених Штатах, Канаді, Бразилії, Китаї та інших аграрно орієнтованих країнах широко застосовуються сучасні методи аналізу даних. Зокрема, великі агрокомпанії та дослідницькі інститути використовують супутникові знімки, метеорологічні дані та машинне навчання для створення моделей прогнозування врожайності. Такі моделі дозволяють враховувати вплив змінних чинників і робити прогнозування більш гнучким та точним.

Методи машинного навчання, такі як регресійні моделі, дерева рішень, нейронні мережі та ансамблеві методи, використовуються для аналізу великих обсягів даних та визначення залежностей між змінними (рис. 1.2).

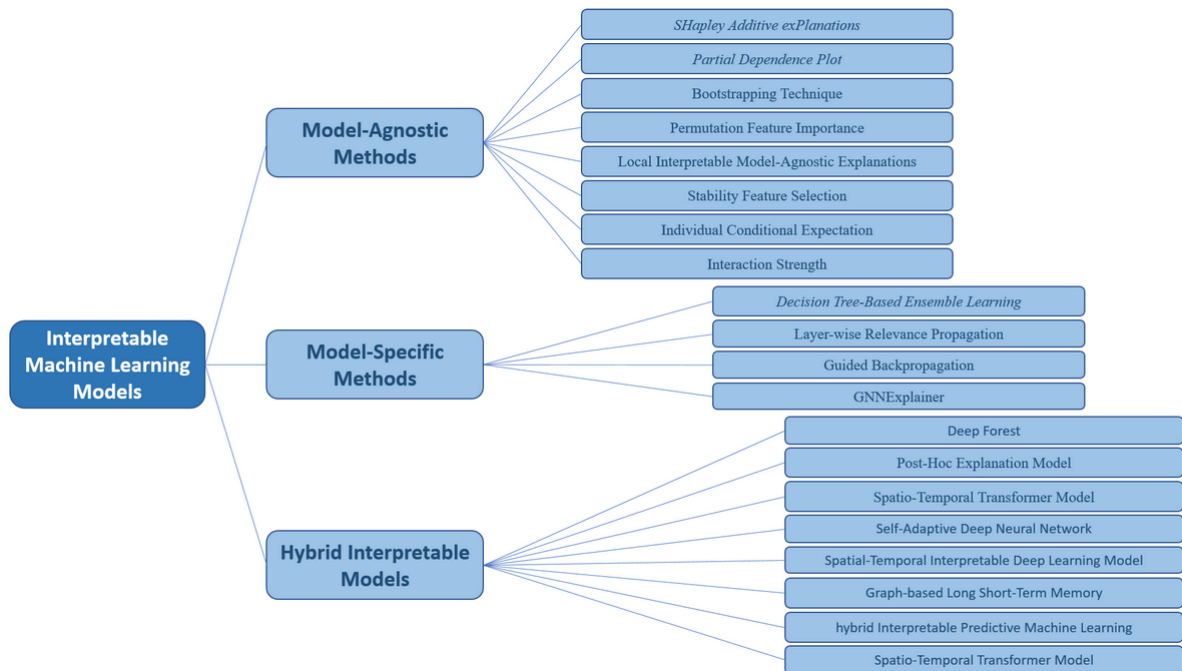


Рисунок 1.2 – Типи інтерпретованих моделей машинного навчання, що використовуються у дослідженнях прогнозування урожайності

Використання цих методів дозволяє передбачати врожайність з урахуванням погодних умов, стану ґрунтів, агротехнічних заходів, біологічних характеристик культур та інших чинників. Успіхи таких методів підтвержені результатами численних досліджень та практичних застосувань у різних регіонах світу.

Попри те, що Україна має значний потенціал у сфері впровадження сучасних методів прогнозування врожайності, розвиток у цьому напрямку поки що поступається світовим лідерам. У розвинених країнах, завдяки розвиненій інфраструктурі та високому рівню інвестицій в технології, застосування машинного навчання та великих даних стало нормою. В Україні ж, хоча й існують окремі приклади використання цих методів, необхідна більша підтримка та фінансування для розширення доступу до технологій та підготовки спеціалістів.

Аналіз стану предметної області показує, що прогнозування врожайності сільськогосподарських культур за допомогою машинного навчання є важливим напрямком для підвищення ефективності аграрного виробництва. В Україні необхідно активніше впроваджувати сучасні методи, використовуючи досвід передових країн та адаптуючи його до національних умов. Це сприятиме зниженню ризиків, забезпеченню стабільності врожаїв та підвищенню конкурентоспроможності країни на світовому ринку.

## **1.2. Доцільність використання машинного навчання для прогнозування врожайності сільськогосподарських культур**

Точні прогнози врожайності сільськогосподарських культур мають вирішальне значення для підтримки сільськогосподарської практики та зменшення негативних наслідків зміни клімату. На відміну від попередніх досліджень, які здебільшого використовують традиційні статистичні методи, наше дослідження включає потужні алгоритми машинного навчання для прогнозування врожайності сільськогосподарських культур за різними сценаріями зміни кліматичних умов [21].

Сільськогосподарські культури є життєво важливим джерелом, що забезпечує їжею приблизно 40% населення світу [8; 4]. Більшість фермерів, особливо в країнах, що розвиваються, покладаються виключно на свої обмежені знання та попередній досвід. Через ці обмежені знання їм важко конкурувати на глобальному рівні та задовольняти зростаючі вимоги [25]. У цьому дослідженні використовуються найсучасніші інструменти для вивчення складної взаємодії між кліматичними умовами та врожайністю сільськогосподарських культур, надаючи більш точні та корисні результати.

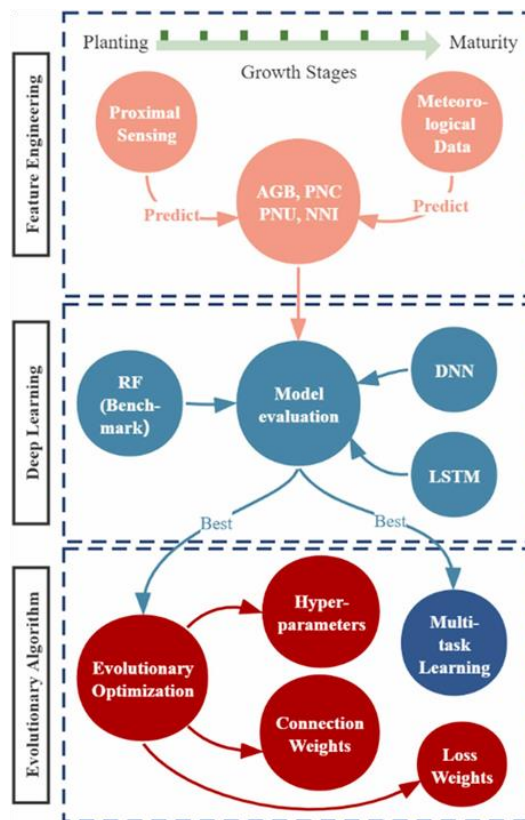


Рисунок 1.3 – Робочий процес системи оцінки статусу азоту пшениці на основі алгоритму еволюції на основі глибокого навчання [21]: AGB: надземна біомаса; PNU: Поглинання азоту рослинами; PNC: концентрація азоту в рослинах; NNI: індекс живлення азотом; РФ: випадковий ліс; DNN: глибока нейронна мережа; LSTM: нейронна мережа довготривалої короткочасної пам'яті.

Екстремальна погода, зміна тенденцій кількості опадів і підвищення температури – все це може зашкодити продовольчій безпеці та сільськогосподарському виробництву. Через невизначеність погоди важко передбачити врожайність сільськогосподарських культур, яка є дуже важливою. Такі речі, як заморозки на етапах цвітіння та наливання зерна, температура вище 30°C і недостатня кількість дощу в основні періоди, можуть призвести до втрат виробництва [1]. Ці проблеми вже існують, і зміна клімату лише погіршить їх. Це ускладнить планування та ефективну обробку посівів.

Клімат, методи управління та особисті стандарти можуть вплинути на зростання врожаю. Дослідники використовують технологічні та статистичні методи, такі як біологічні дані та дистанційне зондування [20; 12; 14], щоб

стежити за виробництвом. Моделі, засновані на процесах, можна використовувати для певних культур, таких як рис, пшениця та кукурудза, щоб допомогти в управлінні ресурсами та розвитку сільського господарства [36]. Ці моделі використовують правильні дані та алгоритми, щоб надавати фермерам інформацію, яка допомагає їм приймати розумні рішення. Статистичні методи також можуть передбачити результати в широкому діапазоні ситуацій, дивлячись на те, як погода та якість землі пов'язані з посівами. Класичні статистичні методи, засновані на математичних моделях, часто використовуються для перевірки ідей на зразках культур. Їхній успіх, однак, залежить від легкого доступу до відповідних сільськогосподарських даних [17].

Глобальні кліматичні моделі (GCM) здебільшого використовуються, щоб припустити, скільки їжі буде вирощено в усьому світі за різних погодних умов та умов управління. Найкращий спосіб дізнатися, як кліматичні зміни вплинуть на зростання врожаю, – це використовувати GCM. GCM використовувалися для моделювання погоди, землі, управління та інших аспектів розвитку сільського господарства з метою прогнозування зростання продовольства та виробництва на глобальному та регіональному рівнях. У різних частинах світу різні погода, земля та способи вирощування. Вчені вважають, що GCM зазвичай мають просторову роздільну здатність  $50 \text{ км} \times 50 \text{ км}$ , що достатньо, щоб отримати гарне уявлення про результати діяльності у сільському господарстві. Завдяки додаванню цих чинників площі зменшення масштабу дає змогу робити більш точні та надійні оцінки врожайності.

Важливо стежити за такими погодними умовами, як температура повітря,  $\text{CO}_2$ , кількість опадів і час росту, коли ви намагаєтеся адаптуватися до зміни клімату. У всьому світі було проведено багато досліджень, щоб з'ясувати, як зміна клімату впливає на врожайність сільськогосподарських культур. Щоб показати, як кліматичні зміни впливають на врожайність, у роботі [26] використано методи підвищення врожайності, які були частиною плану СМІР5. Що стосується роботи [11], то у ній запропоновано одним зі способів зміни

умов завдяки перенесенню найкращого часу для посіву сільськогосподарських культур на більш прохолодні місяці (рис. 1.4).

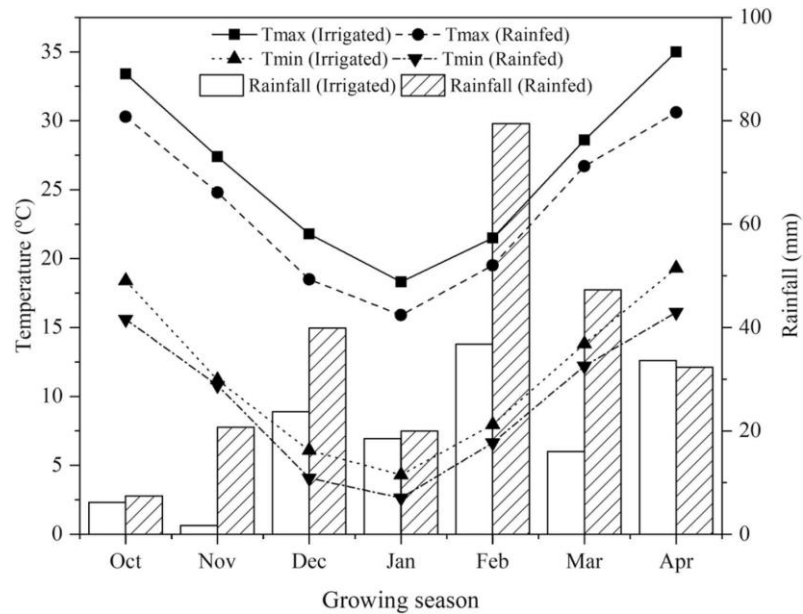


Рисунок 1.4 – Середньомісячні максимальні та мінімальні температури повітря (°C) і кількість опадів (мм) протягом місяців вирощування пшениці дослідження для зрошуваних та реальних умов [11]

Використання машинного навчання для вирішення проблем показало багато перспектив [27; 34]. Вони здатні працювати з даними, які мають багато вимірів, знаходити непрямі зв'язки та бачити складні тенденції [33]. Алгоритми ML знаходять зв'язки між залежними та незалежними факторами, навчаючи їх за допомогою просторово-часового навчання спостереження на дуже великих наборах даних. Методи машинного навчання використовувалися різними способами для ідентифікації культур і прогнозування їх урожайності.

### 1.3. Аналіз використання алгоритмів ML для прогнозування врожайності сільськогосподарських культур

Алгоритми ML допомагають у плануванні та веденні діяльності у сільському господарстві, і вони часто використовуються, щоб припустити,



скільки продовольчих культур дадуть. Для оцінки врожайності використовуються такі моделі, як штучні нейронні мережі (ANN), множинна лінійна регресія (MLR) [6], регресія випадкового лісу (RFR) [18] і XGboost [22]. Ці моделі враховують такі речі, як температура, вологість, швидкість вітру, кількість опадів і атмосферні гази.

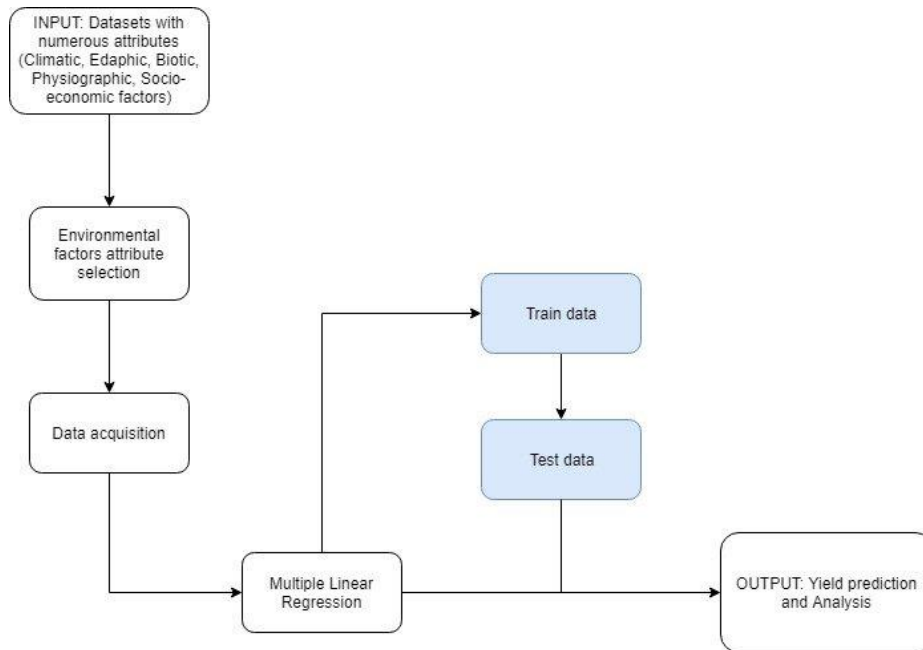


Рисунок 1.5 – Схема дослідження врожайності на основі множинної лінійної регресії (MLR) [6]

Переобладнання є проблемою цих моделей, що робить їх менш корисними, навіть якщо вони добре справляються з проблемами, пов'язаними з кліматом у сільському господарстві. Використання методів ансамблю з декількома моделями робить прогнозування врожайності більш точним, одночасно зменшуючи переобладнання [7; 15]. Ансамблеві методи, такі як метамашинне навчання, можуть підвищити точність прогнозів росту сільськогосподарських культур шляхом поєднання результатів багатьох моделей.

Коли справа доходить до прогнозування зміни клімату, статистичні моделі та старі методи, які базуються на минулих даних, часто не справляються. Ми пропонуємо заповнити цю прогалину, навчивши модель

машинного навчання прогнозувати врожайність сільськогосподарських культур, навіть коли погодні тенденції змінюються.

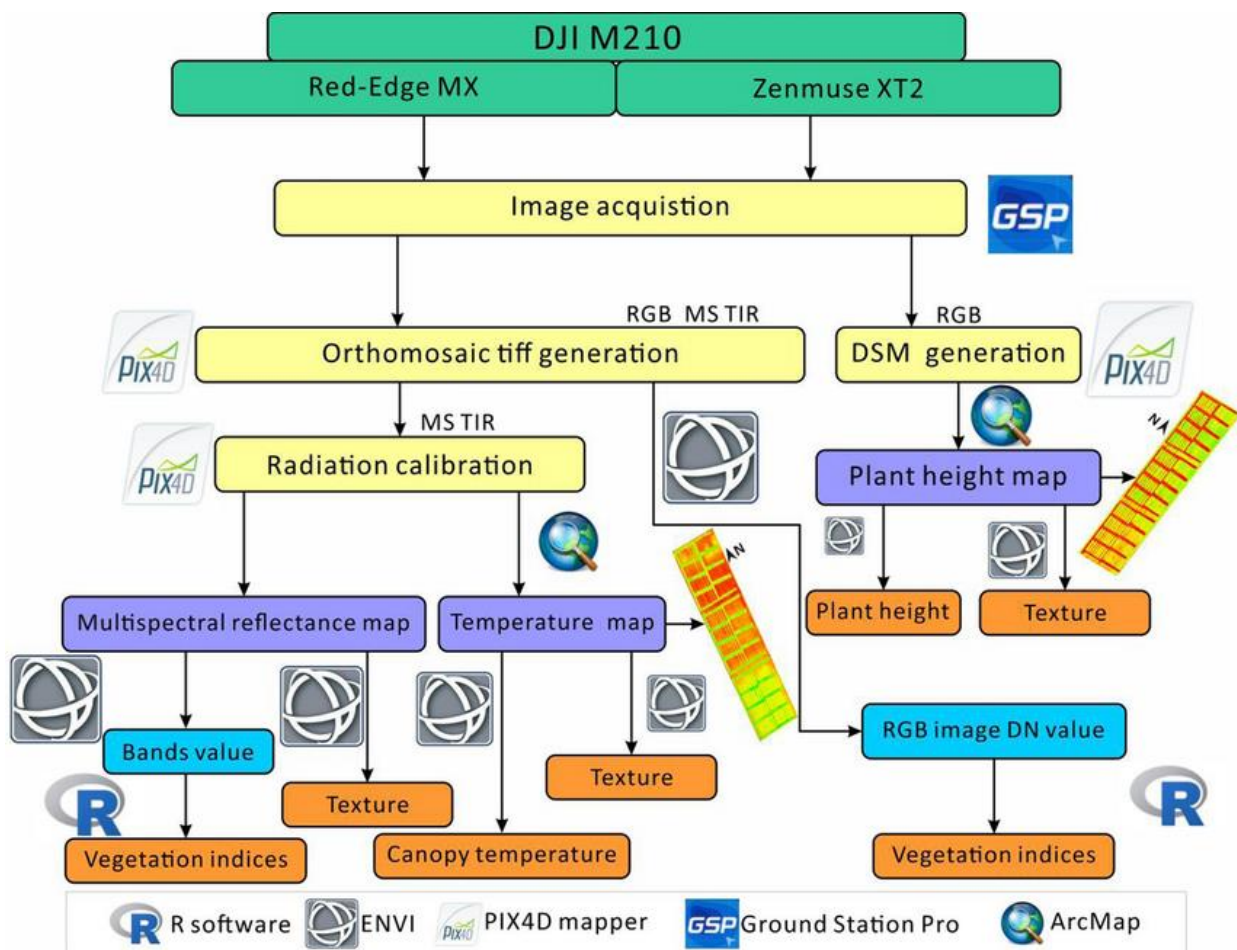


Рисунок 1.6 – Діаграма робочого процесу збору даних, обробки даних і вилучення ознак. Мультиспектральні зображення MS, теплові інфрачервоні зображення TIR, цифрова модель поверхні DSM [7]

Алгоритми машинного навчання можуть використовувати величезні набори даних для вивчення складних кореляцій між змінними та створення точних прогнозів. Запропоноване рішення об'єднує дані метеостанції та майбутні сценарії викидів CO<sub>2</sub> для розробки точної моделі машинного навчання, здатної прогнозувати врожайність сільськогосподарських культур. Враховуючи вплив зміни клімату на врожайність, наша модель дає фермерам можливість адаптувати свою практику до змін.

Очікуваним результатом є підвищення точності прогнозування, підвищення виробництва сільськогосподарських культур та продовольчої

безпеки. Подібні дослідження з використанням аналогічних методологій були проведені в інших країнах для різних культур [35].

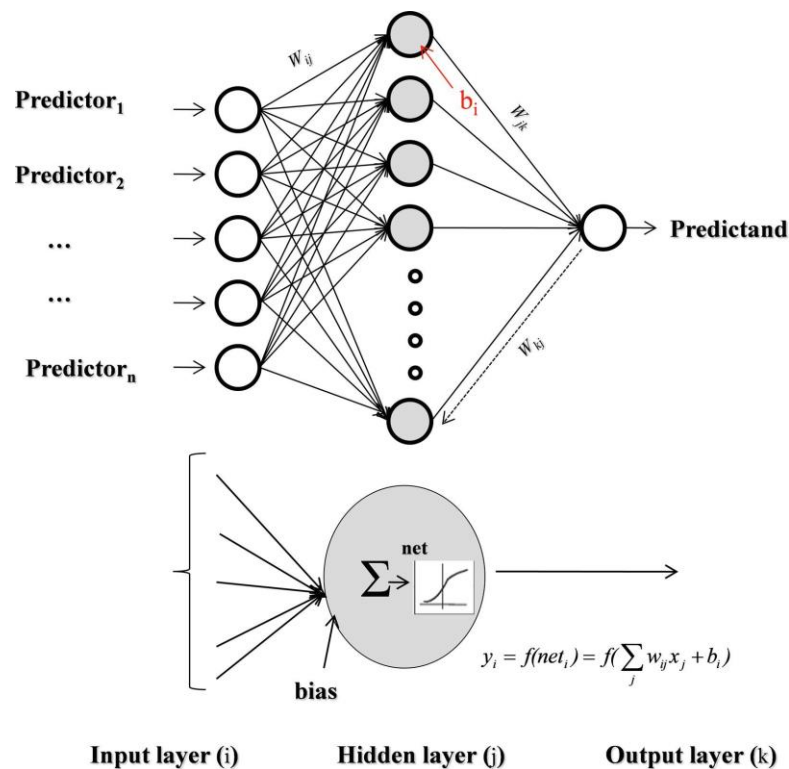


Рисунок 1.7 – Типова структура штучної нейронної мережі для прогнозування виробництва сільськогосподарських культур [35]

Однак вони не використовували ансамблеві моделі, виявляючи прогалину, яку прагне виправити наше дослідження. Наше дослідження вирізняється поєднанням методів ансамблевого моделювання, спрямованих на підвищення надійності та точності прогнозів урожайності.

Декілька досліджень були зосереджені на прогнозуванні виробництва сільськогосподарських культур, особливо в Пакистані, використовуючи різні прогнози, такі як кількість опадів, добрива, температура, трактори та робоча сила. Попередні дослідження підкреслили значну кореляцію між внесенням добрив, методами дистанційного зондування та виходом сільськогосподарських культур [24; 9; 2; 3].

Необхідно провести більш комплексні дослідження, які інтегрують численні моделі машинного навчання спеціально для прогнозування

врожайності сільськогосподарських культур за сценаріями зміни клімату. Хоча окремі моделі досліджувалися, для підвищення точності прогнозу потрібні комплексні методи, які включають переваги багатьох алгоритмів.

На відміну від попередніх досліджень, які в основному поклалися на традиційні статистичні методи або були зосереджені на інших регіонах, ця робота представляє новий спосіб прогнозування врожайності сільськогосподарських культур за сценаріями зміни клімату, який використовує складний ансамбль моделей машинного навчання. Особливістю цього дослідження є інтеграція багатьох передових алгоритмів машинного навчання для оцінки виробництва сільськогосподарських культур на основі кліматичних даних, що передбачає використання історичних кліматичних даних, а також прогнозів із GCM.

Наше дослідження унікально поєднує історичні кліматичні дані, прогнози глобальної кліматичної моделі (GCM) і передові методи машинного навчання. Таким чином, ми надаємо вичерпний і точний прогноз врожайності сільськогосподарських культур спеціально для України. Наш міждисциплінарний підхід сприяє прогнозуванню врожайності сільських культур та пропонує практичні ідеї для розробки стійких до клімату методів землеробства. Ці дослідження спрямовані на підвищення точності прогнозування врожаю шляхом аналізу впливу різних чинників на продуктивність діяльності у сільському господарстві, особливо на виробництво сільськогосподарських культур в Україні.

Основною метою цього дослідження є побудова моделі для прогнозування врожайності сільськогосподарських культур за допомогою ШНМ з використанням метеорологічних даних і даних GCM. Для досягнення поставленої мети були сформульовані такі завдання:

- 1) визначення основні чинники, що впливають на виробництво сільськогосподарських культур;
- 2) моделювання та тестування реакції врожайності сільськогосподарських культур на зміни кількості опадів і температури з

використанням різних методів, таких як стимулювання дерев, ШНМ, регресія випадкового лісу, множинна лінійна регресія та ансамблеві моделі на основі спостережуваної врожайності та кліматичних Даних

3) прогнозування та аналіз потенційного впливу зміни клімату на тенденції посівів сільськогосподарських культур до 2052 року.

#### **1.4. Обґрунтування доцільності прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання**

Прогнозування врожайності сільськогосподарських культур є однією з найважливіших задач у сучасному аграрному секторі. Успішне вирішення цієї задачі сприяє оптимізації виробничих процесів, раціональному використанню ресурсів, зниженню економічних ризиків та підвищенню загальної ефективності аграрного виробництва. Традиційні методи прогнозування часто базуються на статистичних підходах та експертних оцінках, що не завжди забезпечує необхідну точність через складність урахування численних факторів, таких як погодні умови, стан ґрунтів, застосовані агротехнічні заходи та інші параметри.

Методи машинного навчання (ML) пропонують більш гнучкий та точний підхід до прогнозування завдяки можливості аналізу великих обсягів даних, виявлення прихованих взаємозв'язків та адаптації моделей до різних умов. Обґрунтування доцільності використання ML для прогнозування врожайності сільськогосподарських культур ґрунтується на аспектах, які представлено на рис. 1.8.

Однією з основних переваг методів машинного навчання є здатність підвищувати точність прогнозів. Алгоритми ML, такі як нейронні мережі, метод опорних векторів, ансамблеві методи тощо, дозволяють створювати моделі, які враховують численні фактори та їхні взаємозв'язки. У порівнянні з

традиційними методами, ML моделі здатні виявляти складні, нелінійні залежності, що робить їх більш надійними в умовах змінної природи та нестабільних погодних умов.

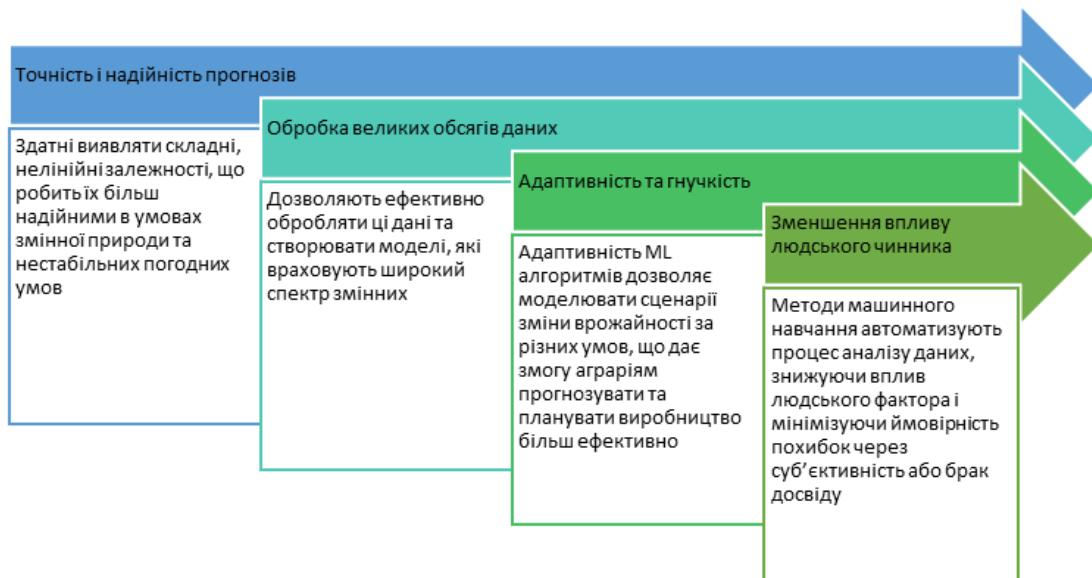


Рисунок 1.8 – Складові доцільності використання ML для прогнозування врожайності сільськогосподарських культур

Сучасні аграрні підприємства накопичують великі масиви даних, включаючи інформацію про погодні умови, типи ґрунтів, добрива, застосовані технології та історичні врожаї. Методи машинного навчання дозволяють ефективно обробляти ці дані та створювати моделі, які враховують широкий спектр змінних. Завдяки цьому аграрії отримують можливість оперативно адаптувати плани виробництва та реагувати на зміни в умовах вирощування.

Традиційні методи прогнозування значною мірою залежать від суб'єктивних експертних оцінок, що може призвести до помилок. Методи машинного навчання автоматизують процес аналізу даних, знижуючи вплив людського фактора і мінімізуючи ймовірність помилок через суб'єктивність або брак досвіду. Це особливо важливо в умовах великого обсягу інформації та складних взаємозв'язків між змінними.

Моделі машинного навчання можуть бути адаптовані до специфічних умов регіону або окремого господарства. Це забезпечує можливість налаштування алгоритмів під конкретні типи культур, кліматичні умови та

особливості технологічних процесів. Адаптивність ML алгоритмів дозволяє моделювати сценарії зміни врожайності за різних умов, що дає змогу аграріям прогнозувати та планувати виробництво більш ефективно.

Завдяки точним прогнозам врожайності аграрні підприємства можуть планувати використання ресурсів більш ефективно, оптимізуючи витрати на добрива, насіння, зрошення тощо. Це сприяє зниженню витрат і підвищенню прибутковості. Зокрема, аграрії можуть заздалегідь передбачати потребу в додаткових заходах з покращення умов вирощування, що забезпечує максимізацію врожайності.

На світовій арені існує безліч прикладів успішного використання методів машинного навчання для прогнозування врожайності. У країнах з розвиненим аграрним сектором, таких як США, Бразилія, Канада та інші, застосування ML дозволяє значно підвищувати точність прогнозів та адаптуватися до змінних умов. В Україні також існує потенціал для впровадження цих технологій з метою підвищення ефективності виробництва та зміцнення конкурентних позицій на світовому ринку.

Отже, тема кваліфікаційної роботи «Прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання» є доцільним та необхідним кроком у напрямку цифрової трансформації аграрного сектору. Це дозволяє забезпечити точність і надійність прогнозів, зменшити ризики та підвищити ефективність управління виробничими процесами. Адаптація ML технологій до аграрних умов України сприятиме підвищенню врожайності, оптимізації ресурсів і стійкості до змінних зовнішніх факторів.

## РОЗДІЛ 2.

### ВИБІР МОДЕЛЕЙ ТА ПІДГОТОВКА ДАНИХ ДЛЯ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР ІЗ ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

#### 2.1. Схема використовуваних моделей для дослідження прогнозування врожайності сільськогосподарських культур

Для проведення дослідження та аналізу даних використовувалися певні матеріали та методи. Ми застосували нову комбінацію моделей машинного навчання для прогнозування врожайності сільськогосподарських культур. Ми ретельно інтегрували регресію випадкового лісу, посилену регресію дерева, штучні нейронні мережі та модель ансамблю. Крім того, ми попередньо обробили кліматичні дані та дані про врожайність і застосували вдосконалені методи зменшення масштабу для підвищення точності локалізованого прогнозу (рис. 2.1).



Рисунок 2.1 – Схема використовуваних моделей для дослідження прогнозування врожайності сільськогосподарських культур



Спочатку історичні дані про температуру, кількість опадів і врожайність збираються та піддаються попередній обробці. Різні моделі машинного навчання, такі як ANN, MLR, RFR і дерева посилення, оцінюються за допомогою попередньо оброблених даних. Потім створюється модель ансамблю, використовуючи сильні сторони цих окремих моделей.

## **2.2. Аналіз використовуваних моделей прогнозування врожайності сільськогосподарських культур**

### **2.2.1. Множинна лінійна регресія**

Множинні моделі лінійної регресії демонструють зв'язок за допомогою лінійного рівняння. Одна або кілька доступних (або прогнозованих) змінних можуть бути пов'язані з окремою залежною (або реагуючою) змінною, тобто кожна незалежна змінна  $X_i$  пов'язана із залежною змінною  $Y$ . Рівняння множинної регресії приймає вигляд:

$$Y = \alpha + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \eta, \quad (2.1)$$

де  $Y$  – змінна врожайності;  $X_1, X_2, X_3$  – позначають кількість опадів відповідно за  $T_{max}$  та  $T_{min}$ ;  $\gamma_1, \gamma_2, \gamma_3$  – наявність дощу відповідно за  $T_{max}$  та  $T_{min}$ ;  $\eta$  – похибкою спостережуваного значення.

### **2.2.2. Екстремальне градієнтне підсилювання XGboost**

Окремі вчені запропонували техніку XGboost, специфічний підхід до посилення градієнта [19]. На рисунку 2.2 показано одне дерево під час розробки послідовних дерев для подальшої оптимізації мети. Модель поєднує прогнози повільних учнів, щоб створити високий рівень навчання, використовуючи

процедури додаткового навчання, засновані на стимулюванні. Модель використовує паралельні обчислення для підвищення швидкості. Він має відмінні здібності підгонки ансамблевого дерева, що є надзвичайно ефективним (щонайбільше в десять разів швидше, ніж радіочастотне навчання) [36]. Базова форма прогнозу в точці  $t$  подана в такому порядку:

$$f_i^t = \sum_{k=1}^t f_k(X_i) = f_i^{(t-1)} + f_t(x_i), \quad (2.2)$$

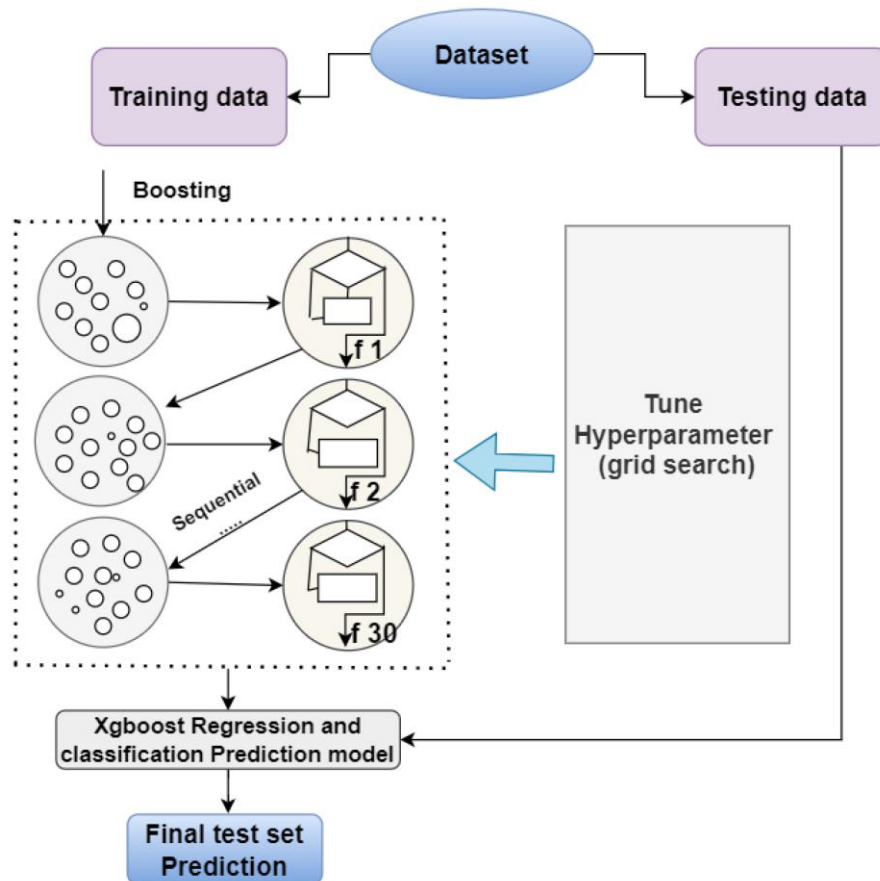


Рисунок 2.2 – Принципова діаграма XGboost

### 2.2.3. Регресія випадкового лісу RFR

Регресія випадкового лісу RFR є популярним методом через його здатність обробляти складні набори даних із нелінійними кореляціями та складними взаємозв'язками ознак. Модель забезпечує надмірну підгонку, а проміжки значень підходять для шумних і неповних даних [5].

RF-алгоритм використовує багато лісів рішень, кожен з яких створюється з використанням даних випадкової вибірки для навчання та коваріації, як показано на рисунку 2.3.

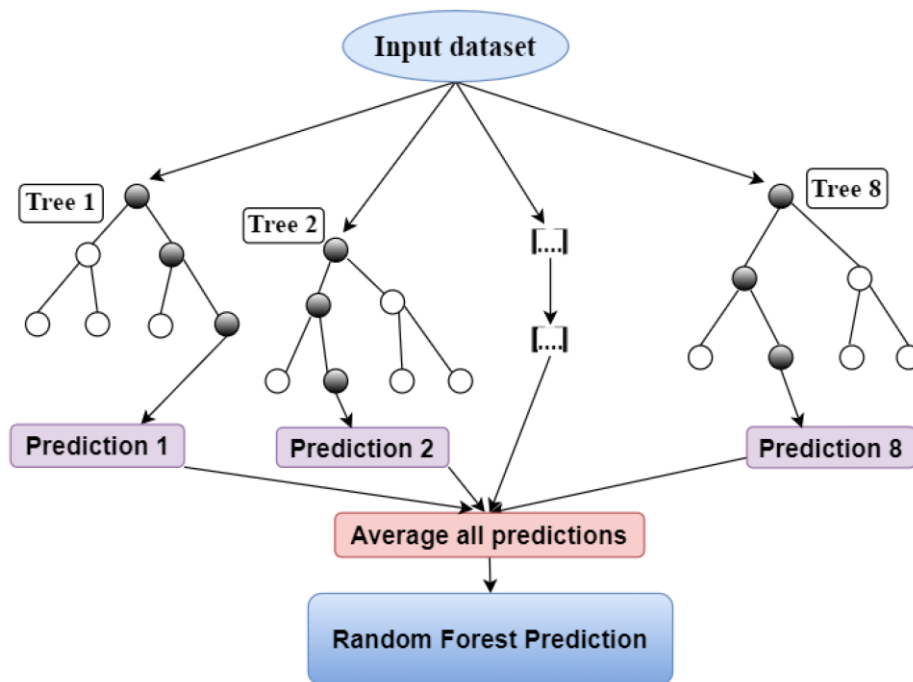


Рисунок 2.3 – Принципова діаграма RFR

Моделі мають перевагу вибору перспективних змінних і тренувальних даних для кожного розбиття, мінімізації переобладнання та підвищення точності прогнозування. Крім того, модель вказує на значимість кожної змінної, яка використовується в прогнозі. Ми використали набір інструментів Scikit-learn, щоб визначити функції RF і важливі функції та оцінити цінність окремих функцій у наших моделях RFR.

Це допомогло нам зрозуміти фундаментальні процеси, які впливають на надійність і продуктивність моделі. Її було оптимізовано за допомогою 20-максимальної глибини з мінімальним розміром розбиття 2. Функція запасу для навчального набору даних, взята випадковим чином із розподілу випадкових векторів  $X$ ,  $Y$  і ансамблю класифікаторів  $h_1(X), h_2(X), \dots, h_k(X)$  полягає в наступному:

$$mg(x, y) = av_k I(h_k(x) = Y) - \max_k av_k I(h_k(X) = j), \quad (2.3)$$

де  $I(h_k(X) = j)$  – індикаторна функція.

Маржа вимірює ступінь, до якого середня кількість голосів у  $X$  та  $Y$  для правильного класу перевищує середню кількість голосів для будь-якого іншого класу.

#### 2.2.4. Модель ансамблю

У роботі [15] порівняли традиційні моделі ML з вдосконаленими моделями ансамблю. Ансамблеві моделі є дуже точними, оскільки вони поєднують численні різновиди одного методу для агрегування прогнозів багатьох базових учнів. У цьому дослідженні ми використовували три моделі для розробки ансамблевої моделі: регресію випадкового лісу, багат шаровий перцептрон (ANN) і підсилювальне дерево. Етапи алгоритму моделі ансамблю для діаграми системи виробництва сільськогосподарських культур показані на рисунку 2.4.

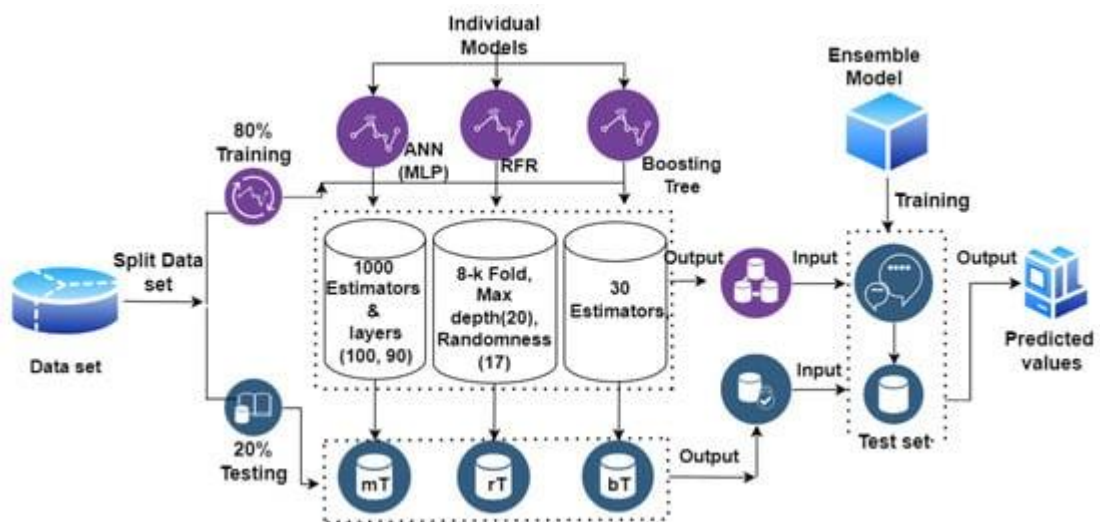


Рисунок 2.4 – Каркасна діаграма моделі ансамблю

Багат шарова модель перцептрона навчається за допомогою 1000 оцінок, використовуючи вхідний і вихідний шари (100, 90). RF-модель навчається зі зміненими параметрами на восьми згортках (за винятком згортки

1), і прогноуються результати згортки 1. Повторення цієї техніки вісім разів генерує прогнозовані значення в навчальному наборі. Для RF ми встановили максимальну глибину 20 і випадковість 17. Модель Boosting Tree навчена за допомогою 30 оцінювачів і випадковість 24.

Навчальний пакет тепер містить три набори очікуваних значень, об'єднаних для створення нового навчального набору. Навчена модель RF, ANN(MLP) і дерева підвищення застосовуються для прогнозування трьох наборів прогнозованих значень для тестового набору (містить N3 зразків). Нарешті, ці три навчені моделі навчають і перевіряють модель ансамблю вторинного рівня. Тестовий зразок подається в підготовлену ансамблеву модель, яка дає кінцевий прогнозований результат. Цей комплексний підхід покращує точність прогнозів і узагальнення, усуваючи обмеження окремих методів машинного навчання в прогнозуванні врожаю.

### 2.2.5. Градієнтне підсилювання Gradient Boosting

Градієнтне підсилювання (Gradient Boosting) – це потужна техніка ансамблевого навчання, яка будує прогноуючу модель поетапно, використовуючи слабкі моделі, зазвичай дерева рішень. Основна ідея полягає у побудові сильної моделі шляхом послідовного додавання нових моделей, кожна з яких коригує помилки попередньої.

Градієнтне підсилювання починається з початкового постійного прогнозу:

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, \gamma), \quad (2.4)$$

де  $L$  – функція втрат (наприклад, середньоквадратична помилка для регресії);  $y_i$  – справжні значення цільової змінної;  $\gamma$  – константа, що мінімізує функцію втрат.

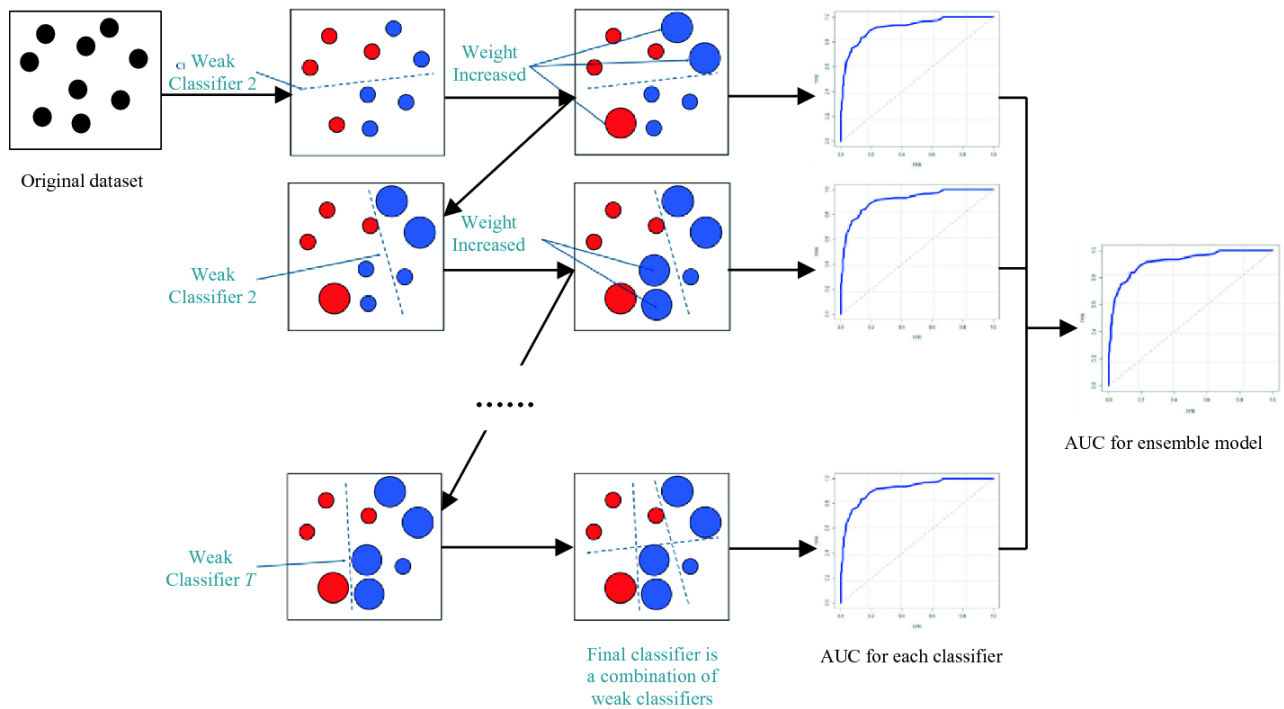


Рисунок 2.4 – Схема градієнтного підсилювання Gradient Boosting

Градієнтне підсилювання будує адитивну модель поетапно:

$$F_m(x) = F_{m-1}(x) + \rho_m h_m(x), \quad (2.5)$$

де  $F_{m-1}(x)$  – прогноз на попередньому етапі.

$h_m(x)$  – новий слабкий учень, доданий на етапі  $m$ .

$\rho_m$  – коефіцієнт навчання, що масштабує внесок  $h_m(x)$ .

Щоб знайти найкращого слабого учня  $h_m(x)$ , градієнтне підсилювання використовує градієнтний спуск для мінімізації втрат:

$$\tilde{y}_i^{(m)} = - \left[ \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)} \right], \quad (2.6)$$

де  $\tilde{y}_i^{(m)}$  – це псевдорезидуали або негативні градієнти функції втрат щодо поточного прогнозу.

Слабкий учень  $h_m(x)$  навчається для прогнозування цих псевдорезидуалів:

$$h_m(x) \approx \tilde{y}_i^{(m)}, \quad (2.7)$$

Адитивна модель оновлюється за допомогою:

$$F_m(x) = F_{m-1}(x) + \rho_m h_m(x), \quad (2.8)$$

Оптимальне значення  $\rho_m$  визначається як:

$$\rho_m = \arg \min_{\rho} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \rho h_m(x_i)), \quad (2.9)$$

Гradientне підсилювання ефективно поєднує простоту дерев рішень із потужністю ансамблевого навчання, створюючи алгоритм, здатний обробляти складні закономірності даних.

### 2.3. Метрики оцінювання

Метрики оцінювання є важливими для аналізу точності та ефективності моделей машинного навчання. У цьому розділі ми розглянемо основні метрики, які використовувалися для оцінки моделей у наведеному коді: середньоквадратична помилка (RMSE) та коефіцієнт детермінації (Accuracy).

Середньоквадратична помилка (Root Mean Squared Error, RMSE) – це метрика, яка вимірює середню відстань між прогнозованими значеннями та фактичними значеннями. RMSE обчислюється як квадратний корінь середньої квадратичної помилки:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (2.10)$$

де  $y_i$  – фактичне значення цільової змінної для  $i$ -го зразка;  $\hat{y}_i$  – прогнозоване значення для  $i$ -го зразка;  $N$  – кількість зразків.

Чим менше значення RMSE, тим точнішою є модель.

Коефіцієнт детермінації ( $R^2$ ) – це метрика, яка показує, яка частка варіації цільової змінної пояснюється моделлю. Він визначається як:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (2.11)$$

де  $\bar{y}$  – середнє значення цільової змінної.

Значення  $R^2$  може змінюватися від 0 до 1:

- ✓  $R^2=1$  вказує на ідеальне співпадіння моделі з даними.
- ✓  $R^2=0$  означає, що модель не пояснює жодної варіації цільової змінної.

RMSE використовується для оцінки абсолютної похибки моделі. Чим менше значення RMSE, тим модель ближча до фактичних значень.

$R^2$  використовується для оцінки відносної точності моделі. Значення, близьке до 1, вказує на те, що модель добре пояснює варіацію даних.

Комбінація цих метрик дозволяє отримати всебічну оцінку продуктивності моделі, забезпечуючи як абсолютну, так і відносну інформацію про точність прогнозу.

## **2.5. Підготовка даних для прогнозування врожайності сільськогосподарських культур**

Прогнозування врожайності сільськогосподарських культур є особливим завданням, яке потребує обробки підготовки вхідних даних. Дані, що використовуються в моделі, повинні бути не лише точними, але й підготовленими таким чином, щоб максимально розрізнити взаємозв'язки між кліматичними умовами, агротехнічними показниками та остаточною врожайністю.

Набір даних для цього дослідження зібраний з двох надійних джерел: 1) Продовольча та сільськогосподарська організація ООН (FAO), яка надає дані щодо врожайності різних культур та кількості використаних пестицидів; 2) Світовий банк, з бази якого отримані кліматичні дані, зокрема середньорічна температура та середня кількість опадів на рік для кожної країни.

Дані FAO використовують для аналізу агротехнічних чинників, таких як врожайність різних культур (наприклад, кукурудзи, пшениці, рису) та кількості



пестицидів, які застосовуються для захисту рослин від хвороб та шкідників. З іншого боку, кліматичні чинники, такі як середньорічна температура та середньорічна кількість опадів, були отримані зі даних Світового банку. Ці дані є більшими для оцінки того, як зміни в кліматичних умовах можуть вплинути на продуктивність культури.

Отриманий набір даних містить наступні колонки:

- 1) Area – країна, для якої надані дані (наприклад, Албанія);
- 2) Item – тип сільськогосподарської культури (наприклад, кукурудза, картопля, рис);
- 3) Year – рік, до якого належать показники;
- 4) hg/ha\_yield – врожайність у гектограмах з гектара;
- 5) average\_rain\_fall\_mm\_per\_year – середньорічна кількість опадів, виміряна в міліметрах;
- 6) pesticides\_tonnes – кількість використаних пестицидів, виміряна в тонах;
- 7) avg\_temp – середня температура за рік у градусах Цельсія.

Ці дані дозволяють аналізувати вплив кліматичних умов (кількість опадів і температура), використання пестицидів і культур на врожайність сільськогосподарських культур.

У першу чергу виконувалася очищення кожного з набору даних. У даних про врожайність культури та кількість пестицидів були відсутні значення та помилкові записи, що допомогло забезпечити цілісність і надійність даних. Далі дані кліматичних показників також були очищені від неповних записів.

Після очищення кожен набір даних було об'єднано в один єдиний масив, який містить всю необхідну інформацію про агротехнічні та кліматичні фактори. Отриманий масив даних, який зберігається у файлі Yield\_df.csv, містить такі колонки: країна, тип культури, середня кількість опадів (мм/рік), середня температура (°C) та застосовані пестициди (тон) (табл. 2.1).

Таблиця 2.1 – Набір даних для прогнозування врожайності  
сільськогосподарських культур

| Unnamed: 0 | Area  | Item           | Year        | hg/ha_yield | average_rain_fall_mm_per_year | pesticides_tonnes | avg_temp |       |
|------------|-------|----------------|-------------|-------------|-------------------------------|-------------------|----------|-------|
| 0          | 0     | Albania        | Maize       | 1990        | 36613                         | 1485.0            | 121.00   | 16.37 |
| 1          | 1     | Albania        | Potatoes    | 1990        | 66667                         | 1485.0            | 121.00   | 16.37 |
| 2          | 2     | Albania        | Rice, paddy | 1990        | 23333                         | 1485.0            | 121.00   | 16.37 |
| 3          | 3     | Albania        | Sorghum     | 1990        | 12500                         | 1485.0            | 121.00   | 16.37 |
| 4          | 4     | Albania        | Soybeans    | 1990        | 7000                          | 1485.0            | 121.00   | 16.37 |
| ...        | ...   | ...            | ...         | ...         | ...                           | ...               | ...      | ...   |
| 27708      | 27708 | United Kingdom | Wheat       | 2013        | 73838                         | 1220.0            | 17673.46 | 8.76  |
| 27709      | 27709 | United Kingdom | Wheat       | 2013        | 73838                         | 1220.0            | 17673.46 | 9.57  |
| 27710      | 27710 | United Kingdom | Wheat       | 2013        | 73838                         | 1220.0            | 17673.46 | 9.57  |
| 27711      | 27711 | United Kingdom | Wheat       | 2013        | 73838                         | 1220.0            | 17673.46 | 8.39  |
| 27712      | 27712 | United Kingdom | Wheat       | 2013        | 73838                         | 1220.0            | 17673.46 | 9.91  |

3591 rows × 8 columns

Ця таблиця є ключовою для аналізу та використання як основа для подальшого моделювання.

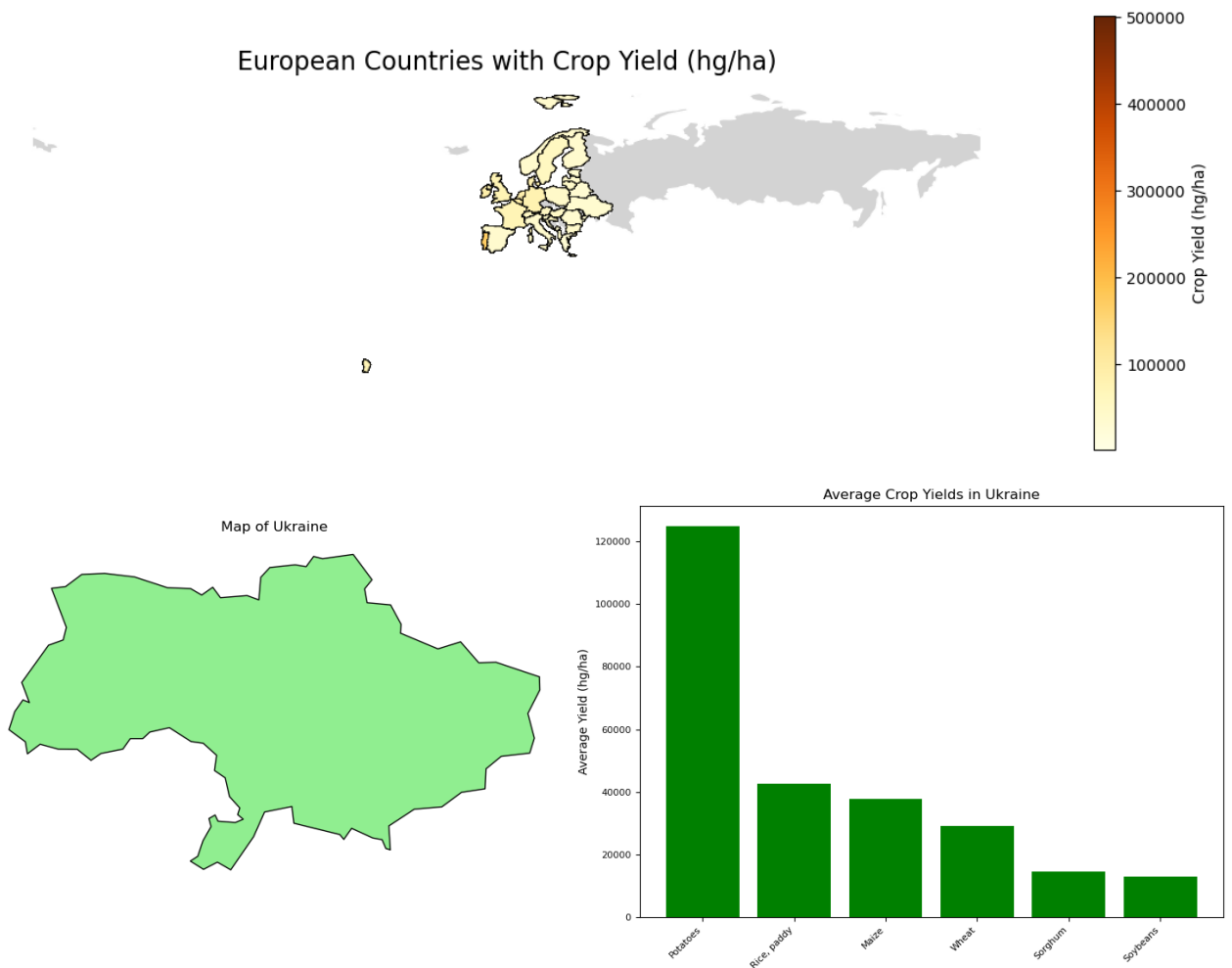


Рисунок 2.5 – Карта із відображення країн із яких було взято дані

Нами виконана візуалізація географічних даних включаючи показники врожайності для різних країн Європи, було створено карту, яка ілюструє розподіл культур та їх врожайність по країнах. На рис. 2.5 представлена карта Європи, що включає країни, для яких були доступні дані, зокрема детально відображена Україна. Це дає можливість проаналізувати просторовий розподіл даних та виявити можливості закономірності, наприклад, вплив кліматичних умов на врожайність певних культур у конкретних регіонах.

Щоб зрозуміти зв'язки між змінами, було виконано кореляційний аналіз між високими показниками, включаючи врожайність, кількість опадів, середню температуру та кількість використаних пестицидів. Нами написано код та створено гістограму, яка ілюструє топ-10 країн за кількістю записів у наборі даних (рис. 2.6).

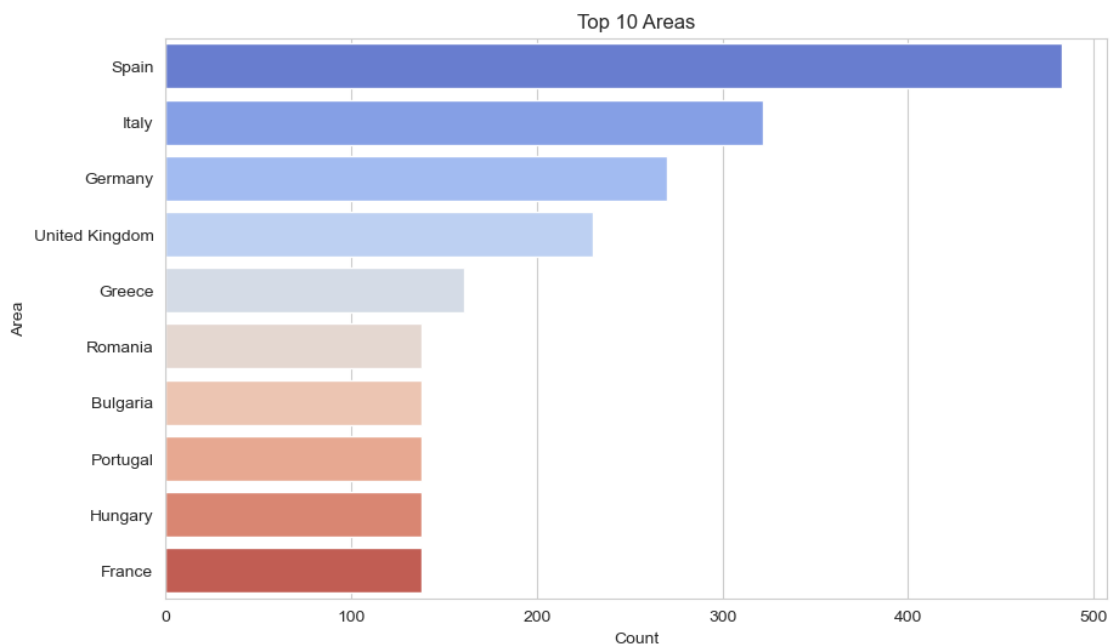


Рисунок 2.6 – Гістограма із топ-10 країн за кількістю записів у наборі даних

Встановлено, що найбільше записів наявно для Іспанії – понад 480 од.; для Італії – понад 320 од.; для Іспанії – понад 270 од.

Нами побудовано розподіл врожайності сільськогосподарських культур за даними стовпця `hg/ha_yield` (рис. 2.7).

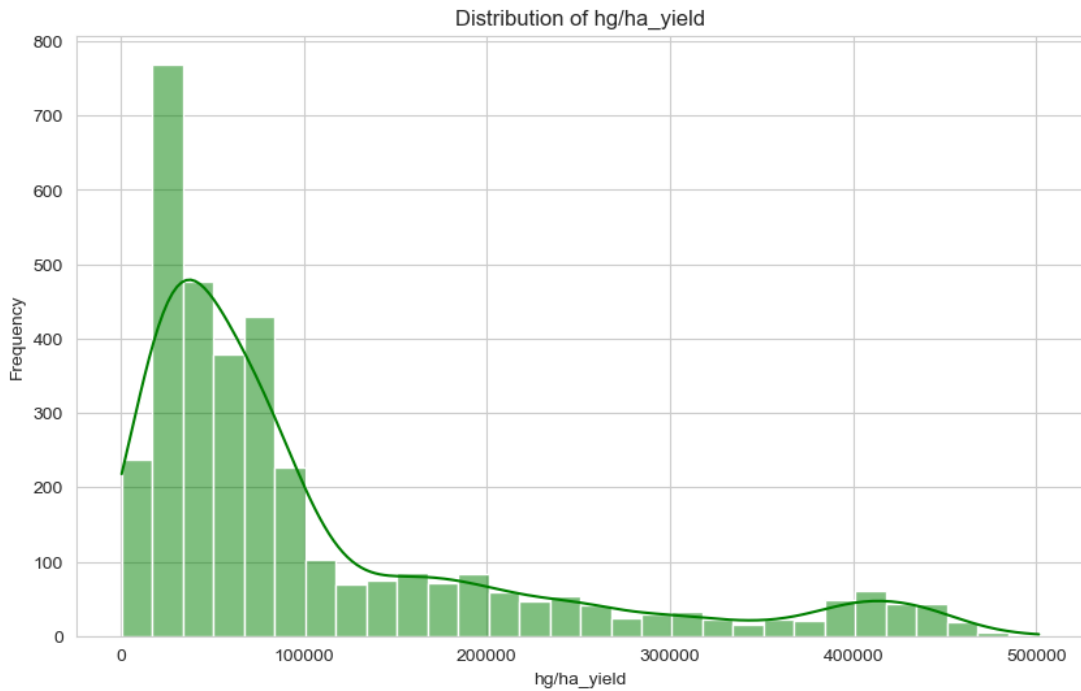


Рисунок 2.7 – Розподіл урожайності сільськогосподарських культур

Для побудови розподілу загалом використано 3591 записів щодо врожайності сільськогосподарських, що вказує на достатню вибірку для аналізу. Середній рівень врожайності становить 106640 гектограм з гектара. Це значення відображає загальну продуктивність культур у вибірці. Стандартне відхилення складає 111930 гектограм з гектара, що вказує на значне коливання врожайності в різних спостереженнях. Це свідчить про широкий діапазон продуктивності, що може бути зумовлено різними факторами, такими як тип культури, кліматичні умови, застосування агротехнічних заходів тощо.

Мінімальне значення врожайності становить 578 гектограм з гектара, що вказує на найнижчий рівень продуктивності в вибірці. Максимальне значення досягає 501412 гектограм з гектара, що відображає найвищий рівень продуктивності у спостереженнях. Це значення може бути пов'язане з особливо високою врожайністю для певних культур або в окремих країнах.

Чверть даних має врожайність, нижчу за 30992 гектограм з гектара. Медіана, або середнє значення, дорівнює 64256 гектограм з гектара. Це вказує, що половина спостережень має врожайність нижчу за це значення, а інша половина – вищу. Три чверті даних мають врожайність, нижчу за 135346.5

гектограм з гектара. Це свідчить, що більшість спостережень зосереджені в діапазоні між 578 та 135346.5 гектограм з гектара, а значення, що перевищують цей рівень, є менш поширеними.

Розподіл врожайності характеризується високою середньою продуктивністю, але значним діапазоном значень. Високе стандартне відхилення та велика різниця між мінімальним і максимальним значеннями свідчать про значні коливання у врожайності сільськогосподарських культур, що ймовірно, пов'язані з типами культур та умовами вирощування в різних регіонах і роках.

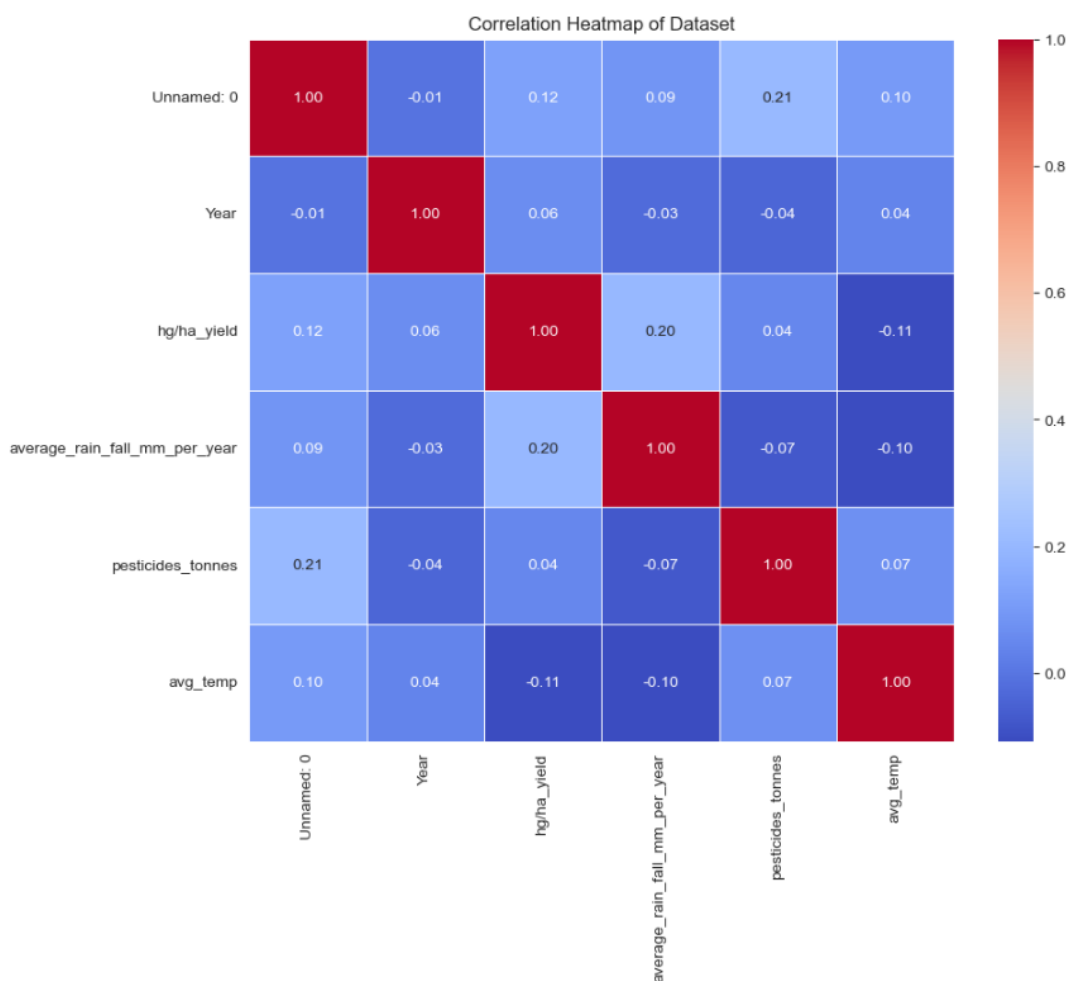


Рисунок 2.8 – Кореляційна теплова карта, яка показує зв'язок між числовими змінними в наборі даних.

На основі аналізу побудовано кореляційну матрицю (рис. 2.8), яка ілюструє взаємозв'язок між змінними. Наприклад, високі коефіцієнти кореляції

між температурою та врожайністю певних культур можуть вказувати на сильний вплив температури на їх продуктивність.

Кореляційна матриця містить коефіцієнти кореляції Пірсона, які показують ступінь зв'язку між парами змінних, з можливими значеннями від -1 до 1. Цей графік дозволяє швидко оцінити взаємозв'язки між змінними в наборі даних, виділяючи ті змінні, які мають сильний позитивний або негативний зв'язок між собою, що є корисним для побудови та інтерпретації моделей прогнозування.

Підготовка даних є важливим кроком у побудові точних моделей прогнозування врожайності сільськогосподарських культур. Об'єднання різних джерел, очищення та візуалізація даних краще зрозуміти їхню структуру та особливості. Таким чином, підготовлено дані для прогнозування врожайності сільськогосподарських культур, що враховують як агротехнічні, так і кліматичні чинники. Це забезпечує належну основу для прогнозування врожайності та є місцем етапу для забезпечення точності моделі.

### РОЗДІЛ 3.

## РЕЗУЛЬТАТИ РОЗРОБКИ МОДЕЛЕЙ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР ІЗ ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

### 3.1. Створення моделей прогнозування врожайності сільськогосподарських культур

Для прогнозування врожайності сільськогосподарських культур було розроблено кілька моделей машинного навчання, включаючи лінійну регресію, модель випадкового лісу, XGBoost, ансамблеву модель та Gradient Boosting. Кожна з моделей має свої особливості, які допомагають забезпечити високу точність прогнозування на основі наявних даних.

Лінійна регресія є простою моделлю, яка намагається описати зв'язок між вхідними змінними та цільовою змінною шляхом побудови лінії найкращого співпадіння. У написаному нами коді для лінійної регресії було використано наступний фрагмент (рис. 3.1).

```
# 1. Multiple Linear Regression
linear_model = LinearRegression()
linear_model.fit(X_train_scaled, y_train)
y_pred_linear = linear_model.predict(X_test_scaled)
rmse_linear = np.sqrt(mean_squared_error(y_test, y_pred_linear))
accuracy_linear = linear_model.score(X_test_scaled, y_test)
```

Рисунок 3.1 – Фрагмент коду створення моделі лінійної регресії для прогнозування врожайності сільськогосподарських культур

Цей фрагмент коду показує, як модель тренується на навчальній вибірці та оцінює свою ефективність на тестовій вибірці. Метрика RMSE (середньоквадратична помилка) та коефіцієнт детермінації (Accuracy) використовуються для оцінки точності прогнозу.

Модель випадкового лісу (Random Forest) складається з ансамблю дерев рішень, які працюють разом для підвищення точності. Ця модель дозволяє захиститися від перенавчання і зазвичай є більш стабільною порівняно з індивідуальними деревами. Код для випадкового лісу представлено на рис. 3.2.

```
# 2. Random Forest Regression
rf_model = RandomForestRegressor(random_state=42)
rf_model.fit(X_train, y_train)
y_pred_rf = rf_model.predict(X_test)
rmse_rf = np.sqrt(mean_squared_error(y_test, y_pred_rf))
accuracy_rf = rf_model.score(X_test, y_test)
```

Рисунок 3.2 – Фрагмент коду створення моделі випадкового лісу (Random Forest) для прогнозування врожайності сільськогосподарських культур

Цей код показує тренування моделі випадкового лісу на навчальних даних, а також оцінку її ефективності на тестовій вибірці.

XGBoost — це вдосконалена модель бустингу, яка добре працює з великими та складними наборами даних. Вона використовує метод градієнтного спуску для мінімізації помилки. Наступний фрагмент коду демонструє, як була реалізована модель XGBoost (рис. 3.3).

```
# 3. XGBoost Regression
xgb_model = XGBRegressor(objective='reg:squarederror', n_estimators=100)
xgb_model.fit(X_train, y_train)
y_pred_xgb = xgb_model.predict(X_test)
rmse_xgb = np.sqrt(mean_squared_error(y_test, y_pred_xgb))
accuracy_xgb = xgb_model.score(X_test, y_test)
```

Рисунок 3.3 – Фрагмент коду створення моделі XGBoost для прогнозування врожайності сільськогосподарських культур

Ця модель XGBoost дозволяє досягти високої точності за рахунок послідовного додавання нових слабких моделей, які намагаються виправити помилки попередніх.



Ансамблева модель об'єднує прогнози декількох інших моделей (у нашому випадку лінійної регресії, випадкового лісу та XGBoost) для отримання кращого результату. Код для ансамблевої моделі представлено на рис. 3.4.

```
# 4. Ensemble Model (Voting Regressor)
ensemble_model = VotingRegressor(['lr', linear_model), ('rf', rf_model), ('xgb', xgb_model)])
ensemble_model.fit(X_train, y_train)
y_pred_ensemble = ensemble_model.predict(X_test)
rmse_ensemble = np.sqrt(mean_squared_error(y_test, y_pred_ensemble))
accuracy_ensemble = ensemble_model.score(X_test, y_test)
```

Рисунок 3.4 – Фрагмент коду створення ансамблевої моделі (Voting Regressor) для прогнозування врожайності сільськогосподарських культур

Ансамблева модель дає змогу отримати більш стабільні результати, оскільки комбінує прогнози кількох моделей.

Градiєнтне підсилювання (Gradient Boosting) є потужним методом ансамблевого навчання, який об'єднує кілька слабких моделей (зазвичай дерев рішень) у сильну модель для досягнення більшої точності. Ідея градієнтного підсилювання полягає в тому, щоб поступово додавати нові моделі, які компенсують помилки попередніх. Кожна наступна модель створюється для виправлення недоліків, виявлених у попередніх кроках. Розглянемо ключові етапи цього методу. Код для створення та навчання моделі Gradient Boosting представлено на рис. 3.5.

```
# 5. Gradient Boosting Regressor
gb_model = GradientBoostingRegressor(n_estimators=200, learning_rate=0.1, random_state=42)
gb_model.fit(X_train_scaled, y_train)
y_pred_gb = gb_model.predict(X_test_scaled)
rmse_gb = np.sqrt(mean_squared_error(y_test, y_pred_gb))
accuracy_gb = gb_model.score(X_test_scaled, y_test)
```

Рисунок 3.5 – Фрагмент коду створення моделі Gradient Boosting для прогнозування врожайності сільськогосподарських культур

Градiєнтне підсилювання дозволяє враховувати складні, нелінійні залежності в даних. Завдяки використанню ансамблю слабких моделей і

поступовому навчанні на залишках, воно забезпечує високу точність навіть на великих наборах даних, роблячи його одним з найефективніших методів машинного навчання.

Після створення та навчання кожної моделі буде проведена оцінка за допомогою метрик RMSE та Accuracy. Це дозволяє порівняти точність та ефективність кожної моделі на основі тестових даних. Графічне представлення результатів також дає змогу побачити, наскільки добре кожна модель прогнозує фактичні значення врожайності.

Таким чином, розробка та тестування різних моделей прогнозування врожайності дозволить отримати результат щодо того, яка з них є найкращою для поставленої задачі.

### **3.2. Результати використання створених моделей для прогнозування врожайності сільськогосподарських культур**

Після розробки і навчання різних моделей прогнозування врожайності сільськогосподарських культур було проведено оцінку їхньої ефективності на тестових даних.

Лінійна регресія показала середній рівень точності. Її простота робить цю модель швидкою та легкою для інтерпретації, але вона не завжди здатна врахувати складні, нелінійні зв'язки між факторами, що впливають на врожайність.

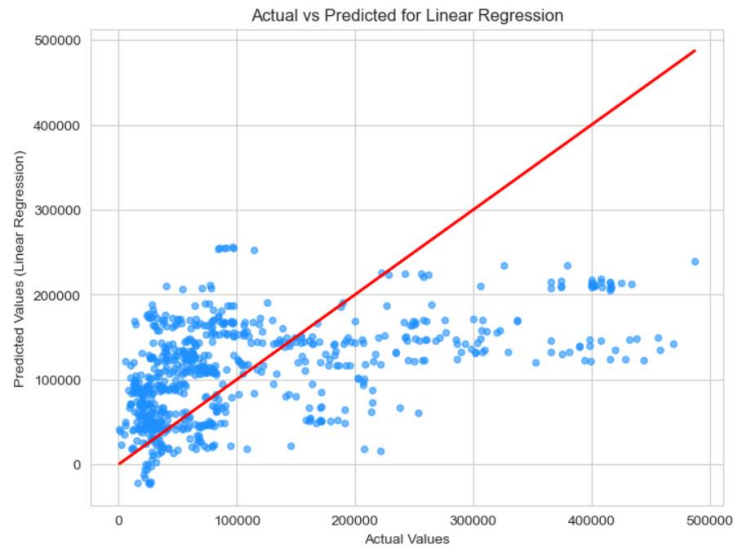


Рисунок 3.6 – Фактичні дані порівняно з прогнозом для моделі лінійної регресії

На рисунку Actual vs Predicted for Linear Regression (рис. 3.6) відображено порівняння фактичних і прогнозованих значень, що демонструє обмеженість моделі в точності прогнозування.

Випадковий ліс продемонстрував значно кращі результати порівняно з лінійною регресією. Завдяки ансамблю дерев рішень модель випадкового лісу може враховувати складні зв'язки та має високу стійкість до шуму в даних. На рисунку Actual vs Predicted for Random Forest (рис. 3.7) можна побачити, що модель забезпечує більш точний збіг між фактичними та прогнозованими значеннями.

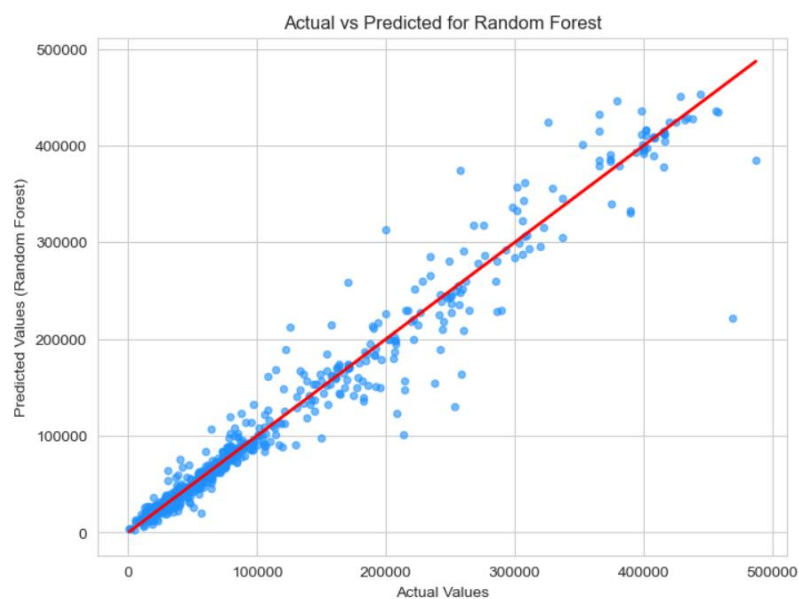


Рисунок 3.7 – Фактичні дані порівняно з прогнозом для моделі Random Forest

XGBoost показав найвищу точність серед усіх розглянутих моделей. Завдяки методам градієнтного підсилювання, ця модель досягла низького значення RMSE, що вказує на здатність ефективно враховувати як лінійні, так і нелінійні зв'язки в даних. На рисунку Actual vs Predicted for XGBoost (рис.3.8) видно, що прогнози дуже близькі до фактичних значень, що робить XGBoost оптимальним вибором для прогнозування врожайності.

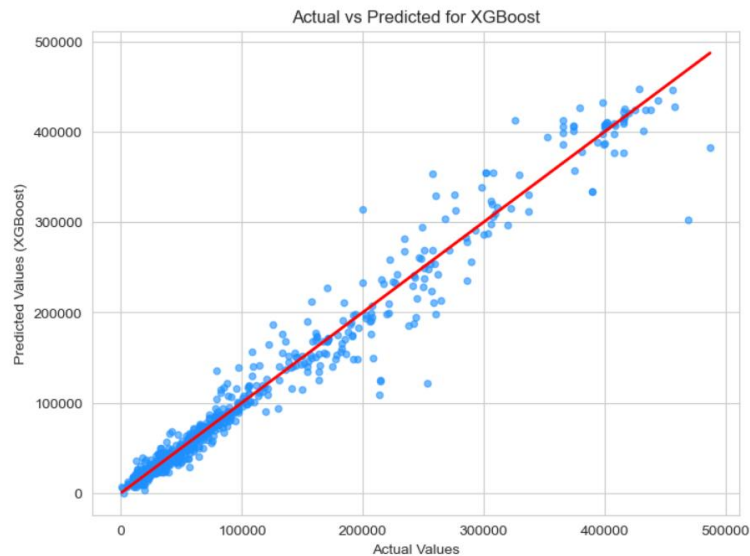


Рисунок 3.8 – Фактичні дані порівняно з прогнозом для моделі XGBoost

Ансамблева модель об'єднала прогнози лінійної регресії, випадкового лісу та XGBoost. Завдяки цьому вона змогла об'єднати переваги всіх моделей, що призвело до хороших результатів. На рисунку Actual vs Predicted for Ensemble Model (рис. 3.9) показано, що ансамблева модель демонструє стабільні прогнози, хоча точність дещо нижча за XGBoost.

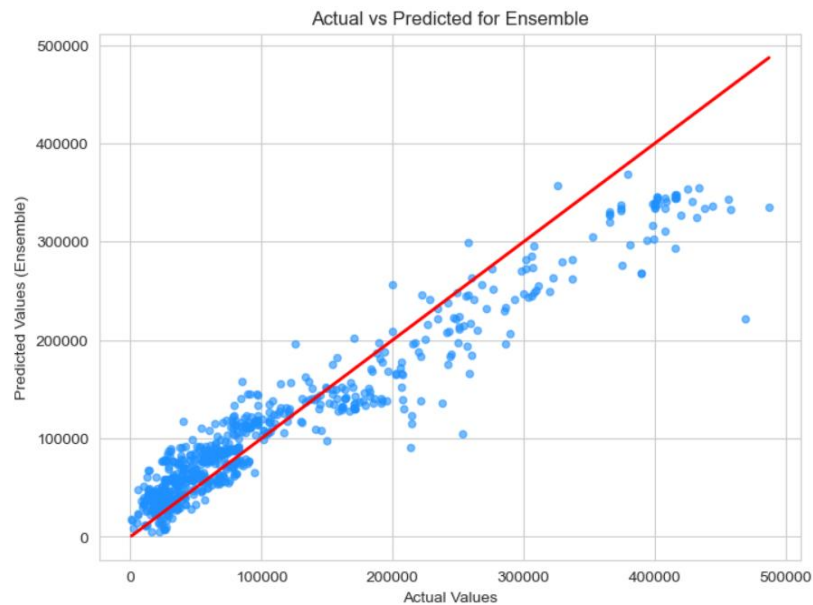


Рисунок 3.9 – Фактичні дані порівняно з прогнозом для моделі ансамблю

Gradient Boosting виявилася менш ефективною порівняно з XGBoost і випадковим лісом. Хоча Gradient Boosting здатна моделювати складні нелінійні залежності, вона вимагала більше обчислювальних ресурсів і часу на тренування. На рисунку Actual vs Predicted for Gradient Boosting (рис. 3.10) видно, що модель має тенденцію до деяких похибок у прогнозуванні, особливо в крайніх значеннях, що може бути результатом обмежень у розмірі даних або структури мережі.

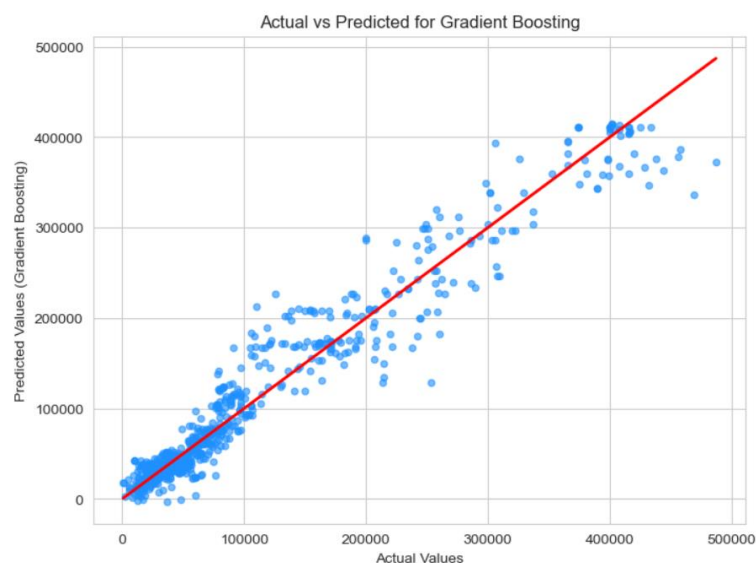


Рисунок 3.10 – Фактичні дані порівняно з прогнозом для моделі Gradient Boosting

Результати підтвердили ефективність сучасних методів машинного навчання для завдання прогнозування врожайності, а XGBoost зарекомендував себе як оптимальний інструмент у цьому дослідженні.

### 3.3. Результати визначення ефективної моделі для прогнозування врожайності сільськогосподарських культур

Для визначення найефективнішої моделі для прогнозування врожайності сільськогосподарських культур було проведено порівняння п'яти моделей на основі метрик середньоквадратичної помилки (RMSE) та коефіцієнта детермінації (Accuracy). Порівняльний аналіз дозволив оцінити кожен модель за точністю та ефективністю, що полегшило вибір оптимальної моделі для цієї задачі (табл. 3.1).

Таблиця 3.1 – Результати визначення кількісних значень RMSE та Accuracy

| Model             | RMSE     | Accuracy |
|-------------------|----------|----------|
| Linear Regression | 94386.06 | 0.2502   |
| Random Forest     | 21420.89 | 0.9614   |
| XGBoost           | 19657.25 | 0.9675   |
| Ensemble          | 35989.04 | 0.8910   |
| Gradient Boosting | 25954.26 | 0.9433   |

Модель лінійної регресії показала найбільше значення RMSE (94386.06), що вказує на високу середню похибку прогнозів. Коефіцієнт детермінації (Accuracy) для цієї моделі становив лише 0.2502, що вказує на низьку здатність до точного прогнозування врожайності.

Модель випадкового лісу Random Forest показала значно нижчий рівень RMSE (21420.89) і високий рівень точності (Accuracy = 0.9614). Це свідчить про стабільні результати та здатність ефективно прогнозувати врожайність.

Найбільш ефективною моделлю за значенням RMSE є XGBoost із значенням 19657.25 та високим коефіцієнтом детермінації 0.9675. Це свідчить про те, що ця модель забезпечує найточніші прогнози серед розглянутих.

Ансамблева модель показала RMSE 35989.04 та Accuracy 0.8910. Хоча ця модель менш точна порівняно з XGBoost і Random Forest, вона демонструє стабільні результати завдяки поєднанню кількох алгоритмів.

Модель Gradient Boosting також показала гарні результати з RMSE 25954.26 і Accuracy 0.9433, але вона дещо поступається XGBoost за точністю.

На побудованій гістограмі RMSE (рис. 3.11) видно, що модель XGBoost має найнижче значення RMSE, що свідчить про її високу точність.

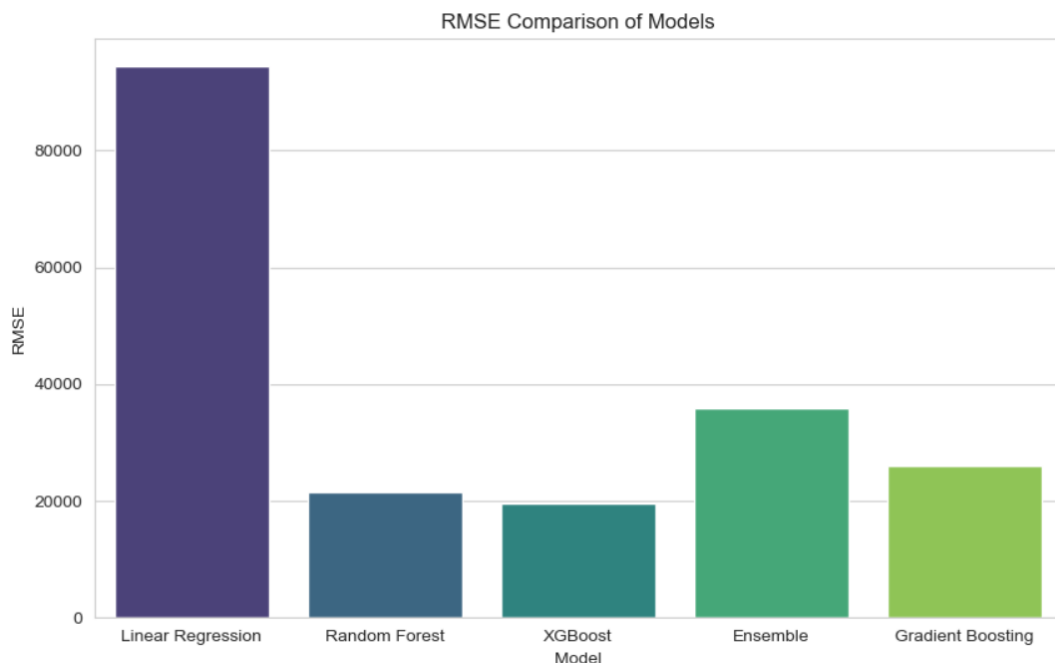


Рисунок 3.11 – Результати визначення середньоквадратичної помилки (RMSE) за використання різних моделей прогнозування врожайності сільськогосподарських культур

На противагу цьому, лінійна регресія має найвищий показник RMSE, що вказує на її обмежену здатність точно прогнозувати врожайність. Моделі випадкового лісу та Gradient Boosting показали середні значення RMSE, що підтверджує їх відносну ефективність у порівнянні з лінійною регресією.

На гістограмі Accuracy (рис. 3.12) видно, що XGBoost і випадковий ліс демонструють найвищі показники точності, наближені до 1. Це підтверджує їх здатність ефективно прогнозувати врожайність. Модель Gradient Boosting також показує високий рівень точності, тоді як лінійна регресія має найнижчий рівень Accuracy, що вказує на її обмежену ефективність для цієї задачі.

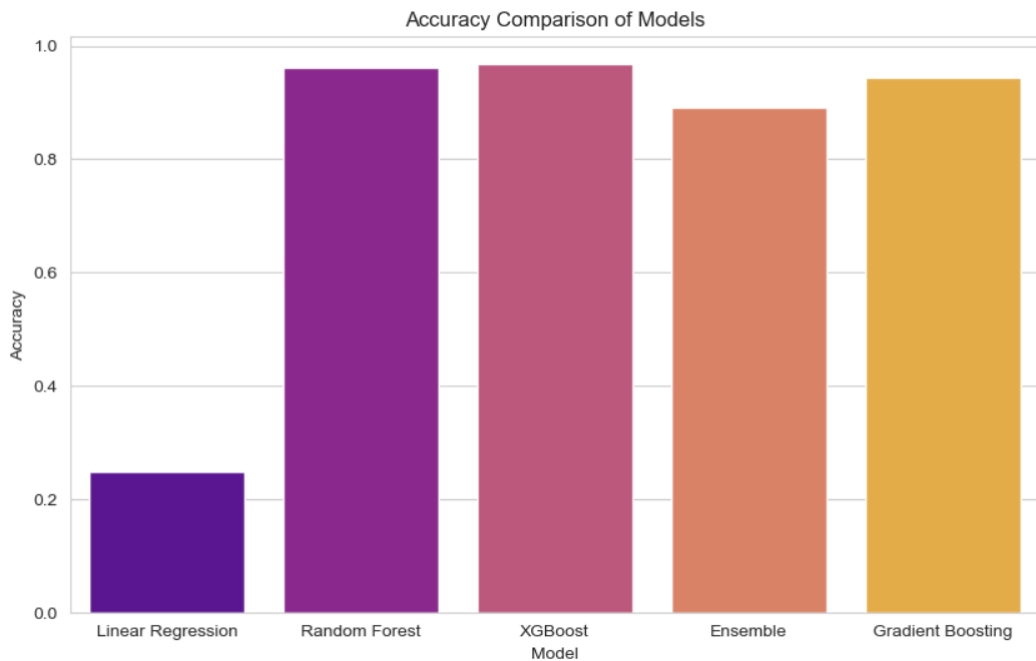


Рисунок 3.12 – Результати визначення коефіцієнта детермінації (Accuracy) за використання різних моделей прогнозування врожайності сільськогосподарських культур

На основі порівняння моделей можна зробити висновок, що модель XGBoost є найефективнішою для прогнозування врожайності сільськогосподарських культур завдяки найнижчому значенню RMSE та найвищому рівню точності (Accuracy). Модель випадкового лісу є гарною альтернативою, а Gradient Boosting також може використовуватись для точного прогнозування, проте поступається XGBoost за деякими показниками. Лінійна регресія виявилася найменш ефективною моделлю в цьому аналізі.



## РОЗДІЛ 4.

### ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

#### 4.1. Небезпечні і шкідливі виробничі чинники під час створення моделей машинного навчання

Процес розробки моделей машинного навчання включає комплекс інтелектуальних, технічних та організаційних заходів, які можуть мати як позитивний, так і негативний вплив на працівників, що забезпечує відповідні завдання. Створення таких моделей вимагає обробки великих масивів даних, використання потужних обчислювальних ресурсів і глибокого занурення в процесі розробки та тестування моделей, що призводить до виникнення небезпечних і шкідливих виробничих чинників. Небезпечні чинники можуть стати причиною травми, а шкідливі – призвести до професійних захворювань, погіршення здоров'я або працездатності працівників.

Розглянемо основні небезпечні та шкідливі чинники при розробці моделей машинного навчання (рис. 4.1).

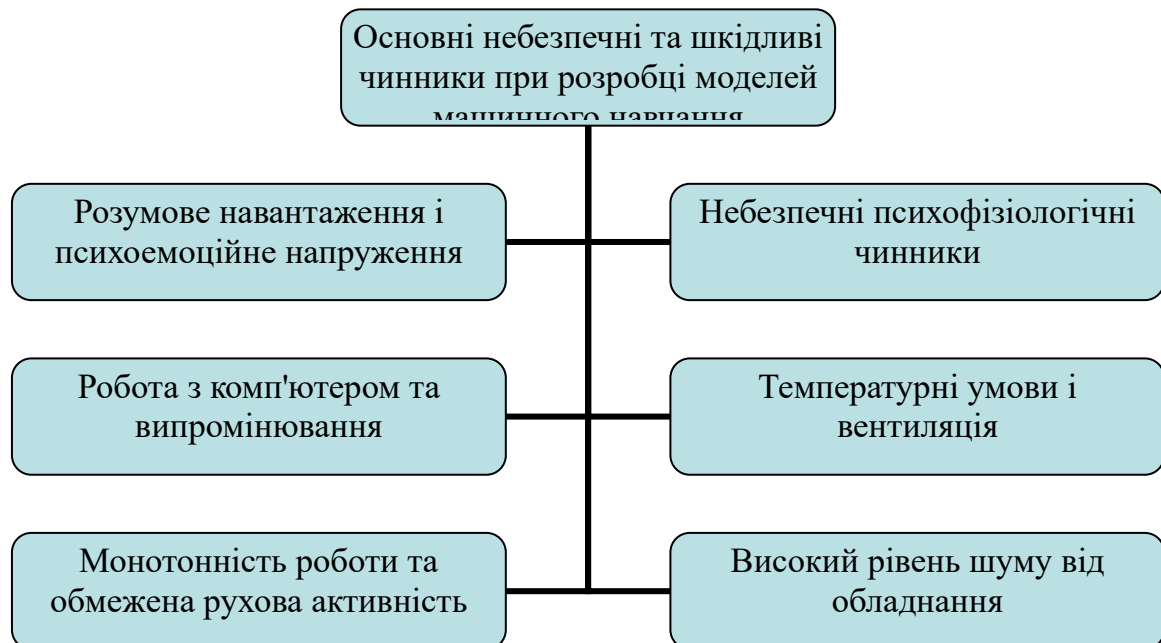


Рисунок 4.1 – Основні небезпечні та шкідливі чинники при розробці моделей машинного навчання

Працівники, зайняті створенням моделей машинного навчання, мають звичайно високе розумове навантаження. Постійне оброблення та аналіз великих обсягів інформації вимагають значної концентрації та призводять до психоемоційного напруження. Це може викликати стрес, швидку швидкість, зниження працездатності та підвищений ризик помилок.

Програми і аналітики машинного навчання часто працюють за комп'ютерами понад 8 годин на день. Така інтенсивна робота може вплинути на зір та загальний стан здоров'я через електромагнітне випромінювання від моніторів та інших електронних пристроїв, а також призвести до розвитку синдрому комп'ютерного зору.

Постійна робота за комп'ютером часто призводить до одноманітності, яка може призвести до зниження рівня уваги та концентрації, збільшити ризик помилок і незначності. крім того, обмеження рухової активності ризик виникнення проблеми із хребтом та порушенням опорно-рухового апарату, що пов'язано з довгою сидячою роботою.

Високий рівень відповідальності за результати роботи, термінові проекти та використання нових технологій можуть спричинити стресові стани. Хронічний стрес може збільшити зниження імунітету, підвищити ризик серцево-судинних захворювань і вплинути на загальне самопочуття.

Під час роботи серверів і комп'ютерного обладнання виділяється значна кількість тепла, що може створити незручності в робочих приміщеннях, якщо не забезпечена належна вентиляція та кондиціонування. Підвищена температура негативного впливу на продуктивність працівників і може призвести до дискомфорту та головного болю.

Сучасне обладнання для обробки великих даних може працювати з високим рівнем шуму, особливо потужні сервери або системи охолодження. Постійний вплив на шум може викликати погіршення, головний біль та зниження концентрації.

Для кращого розуміння показаних небезпек під час створення моделей машинного навчання наведемо таблицю, яка деталізує основні небезпечні та шкідливі виробничі чинники.

Таблиця 4.1 – Небезпечні та шкідливі виробничі чинники

| Чинник                         | Опис впливу   |
|--------------------------------|---|
| Розумове навантаження          | Спричиняє стрес, швидку швидкість, зниження концентрації.                                 |
| Електромагнітне випромінювання | Може вплинути на зір, загальне синдром самопочуття, ушкодження комп'ютерного зору.        |
| Монотонність роботи            | Призвести до зниження уваги, якщо ризик появи помилок та погіршення одноманітності рухів. |
| Обмежена рухова активність     | Викликає проблеми з хребтом, порушення роботи опорно-рухового апарату.                    |
| Хронічний стрес                | Знижує імунітет, ризик серцево-судинних захворювань.                                      |
| Підвищена температура          | Негативно впливає на продуктивність, дискомфорт та головний біль.                         |
| Шум від обладнання             | Викликає емоційність, головний біль, зниження концентрації.                               |

Таким чином, процес створення моделей машинного навчання вимагає дотримання заходів для зниження впливу зазначених діячів.

#### **4.2. Заходи щодо попередження небезпечних і шкідливих виробничих чинників**

Для попередження небезпечних і шкідливих виробничих чинників під час створення моделей машинного навчання необхідно виконати ряд заходів. Забезпечити належне утримання обладнання та робочих місць. Устаткування для машинного навчання та робочі місця повинні регулярно перевірятися та утримуватися в належному стані. Це допоможе запобігти виникненню електротравм, отруень, опіків, травм опорно-рухового апарату та інших травм.

Запровадити заходи щодо захисту працівників від дії фізичних факторів. Працівники, які працюють з обладнанням для машинного навчання, повинні бути забезпечені засобами індивідуального захисту (ЗІЗ). Це допоможе запобігти дії шуму, вібрації, підвищеної температури та запиленості.

Забезпечити раціональну організацію праці. Працівники, які працюють над створенням моделей машинного навчання, повинні дотримуватися норм тривалості робочого часу, робочого тижня та інших норм трудового законодавства. Це допоможе запобігти перевантаженню праці, монотонності праці та стресу.

Створення моделей машинного навчання може бути пов'язане з рядом небезпечних і шкідливих виробничих чинників. Ці чинники можуть призвести до травм, захворювань і навіть смерті працівників. Для попередження небезпечних і шкідливих виробничих чинників необхідно забезпечити належне утримання обладнання та робочих місць, запровадити заходи щодо захисту працівників від дії фізичних факторів та забезпечити раціональну організацію праці.

#### **4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час монтажу інтелектуальної інформаційної системи прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання**

Для побудови логіко-імітаційної моделі процесу, формування і виникнення аварії та травми під час монтажу інтелектуальної інформаційної системи прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання складемо список базових подій. Вони лежатимуть у основі даної моделі. Кожному пункту списку присвоюємо певне значення ймовірності виникнення. Нижче подано сам список:

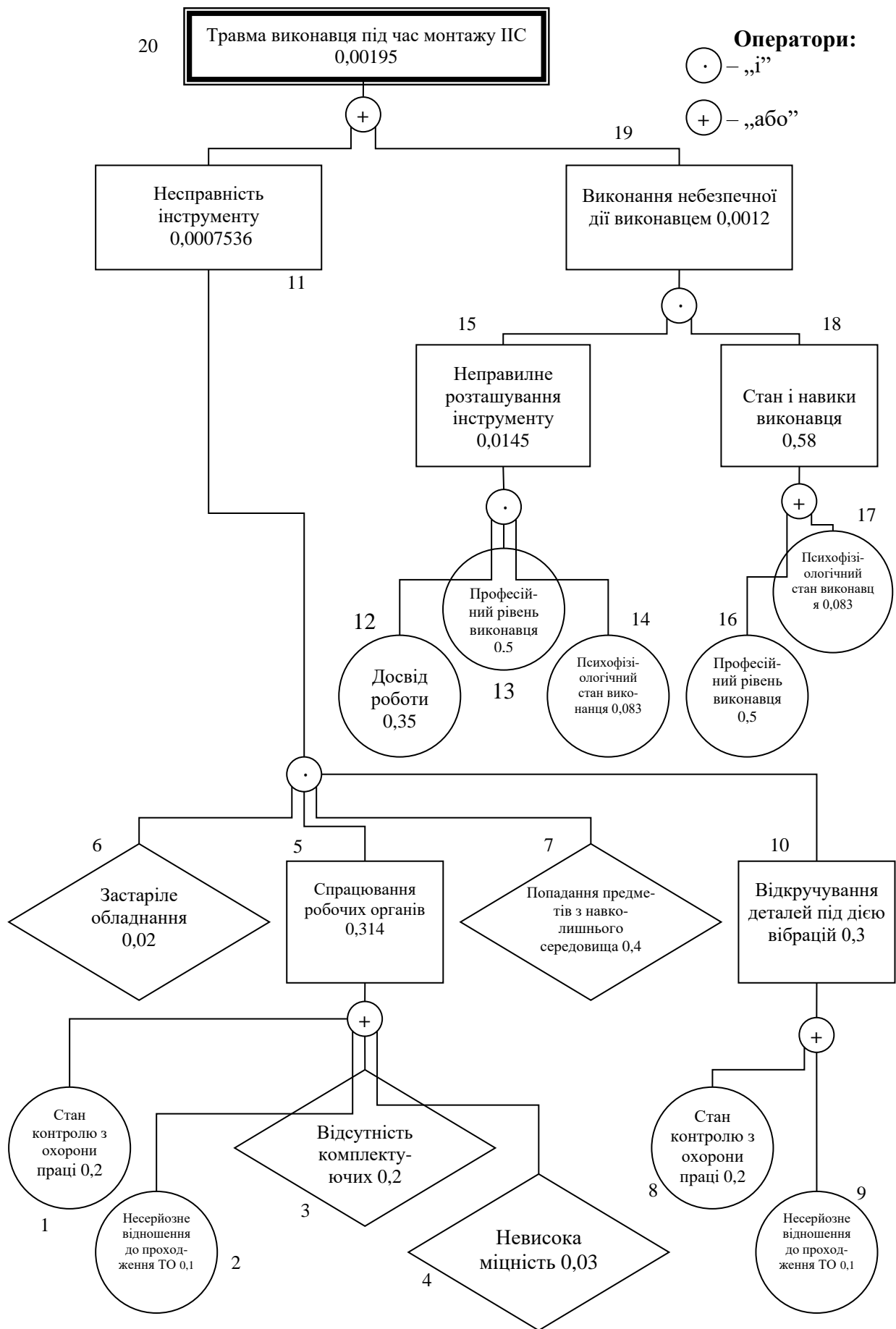


Рис. 4.1. Логіко-імітаційна модель процесу формування та виникнення аварії та травми під час монтажу інтелектуальної інформаційної системи прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання

- |  |                   |
|--|-------------------|
| 1. Стан контролю з охорони праці .....                 | $P_1 = 0,2;$      |
| 2. Несерйозне відношення до проходження ТО інструменту | $P_2 = 0,1;$      |
| 3. Відсутність комплектуючих установки.....            | $P_3 = 0,2;$      |
| 4. Невисока міцність .....                             | $P_4 = 0,03;$     |
| 5. Використання застарілого обладнання.....            | $P_6 = 0,02;$     |
| 6. Попадання сторонніх предметів .....                 | $P_7 = 0,4;$      |
| 7. Досвід роботи виконавця .....                       | $P_{12} = 0,35.$  |
| 8. Професійний рівень виконавця .....                  | $P_{13} = 0,5;$   |
| 9. Психофізіологічний стан виконавця.....              | $P_{14} = 0,083;$ |

На основі даного списку будуюмо матрицю логічних взаємозв'язків між окремими пунктами, графічне представлення якої зображено на рис. 4.1.

Розрахуємо ймовірності виникнення подій, що входять у дану логіко-імітаційну модель процесу монтажу інтелектуальної інформаційної системи прогнозування врожайності сільськогосподарських культур із використанням методів машинного навчання (на прикладі ймовірності отримання травми виконавця).

Ймовірність виникнення події  $P_5$  визначаємо наступним чином:

$$P_5 = 0,2 + 0,1 + 0,2 + 0,003 - 0,2 \cdot 0,1 - 0,2 \cdot 0,03 - 0,2 \cdot 0,03 - 0,1 \cdot 0,2 - 0,1 \cdot 0,03 - 0,2 \cdot 0,03 + 0,2 \cdot 0,1 \cdot 0,2 + 0,1 \cdot 0,2 \cdot 0,03 + 0,2 \cdot 0,1 \cdot 0,2 + 0,2 \cdot 0,1 \cdot 0,03 - 0,2 \cdot 0,1 \cdot 0,2 \cdot 0,03 = 0,314$$

Ймовірність виникнення події  $P_{10}$  визначаємо так:

$$P_{10} = 0,2 + 0,1 = 0,3.$$

Ймовірність виникнення події  $P_{11}$  визначаємо:

$$P_{11} = 0,02 \cdot 0,314 \cdot 0,4 \cdot 0,3 = 0,00075.$$

Ймовірність виникнення події  $P_{15}$  визначаємо наступним чином:

$$P_{15} = 0,35 \cdot 0,5 \cdot 0,083 = 0,0145.$$

Ймовірність події  $P_{18}$ :

$$P_{18} = 0,5 + 0,083 = 0,58.$$

Ймовірність події  $P_{19}$ :

$$P_{19} = 0,0145 \cdot 0,083 = 0,0012.$$

Ймовірність події  $P_{20}$ :

$$P_{20} = 0,00075 + 0,0012 = 0,00195.$$

Ймовірність травми рівна ймовірності виникнення аварії, бо остання можлива лише за умови монтажу системи прогнозування врожайності сільськогосподарських культур людиною.

Логіко-імітаційні моделі аварій і травм допомагають зменшити ймовірність виникнення аварійних та травмонебезпечних ситуацій. Якщо необхідно оцінити рівень небезпеки будь-якого робочого місця, слід уважно вивчити і побудувати логічні моделі можливих небезпечних ситуацій, які охоплюють як стан обладнання і самого робочого місця, так і поведінку працюючого і обчислити ймовірність виникнення травми.

Після аналізу результатів моделювання ймовірність виникнення травми можна звести до дуже малої величини – достатньо зменшити вплив ймовірностей вихідних факторів, які до неї призводять.

#### **4.4. Розробка заходів із забезпечення безпеки під час надзвичайних ситуацій**

Заходи щодо безпеки у надзвичайних ситуаціях – це заходи, які вживаються для запобігання, підготовки до, реагування на та відновлення після надзвичайних ситуацій. Ці заходи спрямовані на захист людей, майна та навколишнього середовища від потенційних збитків.

Розробка заходів щодо безпеки у надзвичайних ситуаціях передбачає кілька етапів. Насамперед виконується оцінка ризиків. На цьому етапі проводиться оцінка потенційних ризиків, які можуть виникнути внаслідок надзвичайних ситуацій. Оцінка ризиків включає в себе аналіз потенційних загроз, ймовірності їх виникнення та можливих наслідків.

Після цього виконується розробка планів. На основі результатів оцінки ризиків розробляються плани заходів щодо безпеки у надзвичайних ситуаціях. Плани заходів включають в себе заходи щодо запобігання, підготовки до, реагування на та відновлення після надзвичайних ситуацій.

Важливим етапом є впровадження планів. Розроблені плани заходів щодо безпеки у надзвичайних ситуаціях повинні бути впроваджені в практику. Це означає, що всі зацікавлені сторони, включаючи персонал, повинні бути проінформовані про плани та навчені їх виконувати.

Завершальним етапом є відстеження та оцінка. Плани заходів щодо безпеки у надзвичайних ситуаціях повинні регулярно відстежуватися та оцінюватися. Це допоможе забезпечити їх ефективність та внести необхідні зміни.

Заходи щодо безпеки у надзвичайних ситуаціях є важливим елементом будь-якої організації. Вони допомагають захистити людей, майно та навколишнє середовище від потенційних збитків, а також зменшити негативні наслідки надзвичайних ситуацій.



## РОЗДІЛ 5.

### ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

Ефективність використання моделі машинного навчання для прогнозування врожайності сільськогосподарських культур можна оцінити за кількома критеріями, включаючи точність моделі, економічний ефект від зменшення втрат урожаю та зниження витрат на агротехнічні заходи.

Економічний ефект від впровадження моделі можна оцінити за допомогою показника збільшення доходу за рахунок точнішого прогнозування врожайності, що дозволяє краще планувати агротехнічні заходи. Формула для розрахунку економічної вигоди за зменшенням втрат урожаю:

$$E = \Delta Y \times P_E, \quad (5.1)$$

де  $E$  – економічний ефект (додатковий дохід);  $\Delta Y$  – зменшення втрат урожаю завдяки поточному прогнозу (в тонах);  $P_E$  – ціна за одиницю продукції (грн/тону).

Припустимо, що точніший прогноз дозволяє зменшити втрати врожаю на певний відсоток середнього рівня втрат. Це можна розрахувати як:

$$\Delta Y = Y \times B_c \times K, \quad (5.2)$$

де  $Y$  – загальна врожайність сільськогосподарської культури;  $B_c$  – середній відсоток втрат врожаю без прогнозування;  $K$  – відсоток зниження втрат завдяки використанню моделі.

Розглянемо приклад, в якому загальна врожайність культури  $Y$  становить 1000 тон. Середній рівень втрати врожаю без прогнозування становить 10%. Використання моделі дозволяє знизити втрати на 20%. Ціна на одиницю продукції становить 8000 грн/тоне.

Проведемо відповідні розрахунки. Насамперед визначаємо зменшення втрат врожаю:

$$\Delta Y = 1000 \times 0,10 \times 0,20 = 20 \text{ тон.}$$

Визначаємо економічний ефект від використання моделі машинного навчання для прогнозування врожайності сільськогосподарських культур:

$$E = 20 \times 8000 = 160000 \text{ грн.}$$

Таким чином, економічний ефект від використання моделі для прогнозування врожайності становить 160000 грн.

Результати виконаних розрахунків представлено у таблиці 5.1.

Таблиця 5.1 – Результати визначення ефективності від використання моделі машинного навчання для прогнозування врожайності сільськогосподарських культур

| Параметр                                | Значення       |
|---|----------------|
| Загальна врожайність $Y$                | 1000 тон.      |
| Середній рівень втрат без прогнозування | 10%.           |
| Коефіцієнти зменшення втрат             | 20%.           |
| Зменшення втрат $\Delta Y$              | 20 тон.        |
| Ціна на тону продукцію $P_E$            | 8000 грн/тону. |
| Економічний ефект $E$                   | 160000 грн.    |

Результати визначення ефективності свідчить, що застосування моделі машинного навчання для прогнозування врожайності дозволяє отримати суттєвий економічний ефект у розмірі 160000 грн. за рахунок зменшення втрат врожаю та поточного планування агротехнічних заходів.

## ВИСНОВКИ І ПРОПОЗИЦІЇ

На сьогоднішній день, завдяки розвитку інформаційних технологій, аграрна галузь має можливість впроваджувати інноваційні підходи до обробки даних. Зокрема, машинне навчання (ML) стало важливим інструментом у моделюванні та прогнозуванні, що забезпечує ефективне використання наявних ресурсів і підвищення продуктивності. Метою даної роботи є розробка підходів до прогнозування врожайності сільськогосподарських культур з використанням методів машинного навчання, що забезпечують підвищення точності прогнозів і сприяють оптимізації процесів планування та управління виробництвом.

Аналіз стану предметної області показує, що прогнозування врожайності сільськогосподарських культур за допомогою машинного навчання є важливим напрямком для підвищення ефективності аграрного виробництва. В Україні необхідно активніше впроваджувати сучасні методи, використовуючи досвід передових країн та адаптуючи його до національних умов. Це сприятиме зниженню ризиків, забезпеченню стабільності врожаїв та підвищенню конкурентоспроможності країни на світовому ринку.

Використання машинного навчання для вирішення задач у різних прикладних сферах показало багато перспектив. Вони здатні працювати з даними, які мають багато вимірів, знаходити непрямі зв'язки та бачити складні тенденції. Алгоритми ML знаходять зв'язки між залежними та незалежними факторами, навчаючи їх за допомогою просторово-часового навчання спостереження на дуже великих наборах даних. Методи машинного навчання використовувалися різними способами для ідентифікації культур і прогнозування їх урожайності.

Виконаний аналіз показав, що алгоритми машинного навчання (ML) є ефективним інструментом для прогнозування врожайності сільськогосподарських культур, забезпечуючи високу точність за рахунок використання великого обсягу даних та складних моделей. Серед найбільш ефективних підходів відзначено використання регресійних моделей та

ансамблевих методів (Random Forest, Gradient Boosting). Вибір оптимального алгоритму залежить від якості даних, їх обсягу та особливостей культури, що прогнозується. Використання ML сприяє підвищенню точності управлінських рішень в аграрному секторі.

Однією з основних переваг методів машинного навчання є здатність підвищувати точність прогнозів. Алгоритми ML, такі як нейронні мережі, метод опорних векторів, ансамблеві методи тощо, дозволяють створювати моделі, які враховують численні фактори та їхні взаємозв'язки. У порівнянні з традиційними методами, ML моделі здатні виявляти складні, нелінійні залежності, що робить їх більш надійними в умовах змінної природи та нестабільних погодних умов.

Для проведення дослідження та аналізу даних пропонується використати різні моделі машинного навчання для прогнозування врожайності сільськогосподарських культур. Зокрема запропоновано дослідити алгоритми регресії випадкового лісу, посиленої регресію дерева та модель ансамблю.

Множинні моделі лінійної регресії демонструють зв'язок за допомогою лінійного рівняння. Одна або кілька доступних (або прогнозованих) змінних можуть бути пов'язані з окремою залежною (або реагуючою) змінною, тобто кожна незалежна змінна пов'язана із залежною змінною .

Модель XGboost має специфічний підхід до посилення градієнта. На рисунку 2.2 показано одне дерево під час розробки послідовних дерев для подальшої оптимізації мети. Модель поєднує прогнози повільних учнів, щоб створити високий рівень навчання, використовуючи процедури додаткового навчання, засновані на стимулюванні. Модель використовує паралельні обчислення для підвищення швидкості.

Регресія випадкового лісу RFR є популярним методом через його здатність обробляти складні набори даних із нелінійними кореляціями та складними взаємозв'язками ознак. Модель забезпечує надмірну підгонку, а проміжки значень підходять для шумних і неповних даних.

Ансамблеві моделі є дуже точними, оскільки вони поєднують численні різновиди одного методу для агрегування прогнозів багатьох базових учнів. Етапи алгоритму моделі ансамблю для діаграми системи виробництва сільськогосподарських культур показані на рисунку 2.4.

Використаний алгоритм градієнтного підсилювання (Gradient Boosting) має потужну техніку ансамблевого навчання, яка будує прогнозуючу модель поетапно, використовуючи слабкі моделі, зазвичай дерева рішень. Основна ідея полягає у побудові сильної моделі шляхом послідовного додавання нових моделей, кожна з яких коригує помилки попередньої.

Критерії оцінювання є важливими для аналізу точності та ефективності моделей машинного навчання. У нашій роботі використано два основні критерії, для оцінки моделей – середньоквадратична помилка (RMSE) та коефіцієнт детермінації (Accuracy).

Прогнозування врожайності сільськогосподарських культур є особливим завданням, яке потребує обробки підготовки вхідних даних. Дані, що використовуються в моделі, повинні бути не лише точними, але й підготовленими таким чином, щоб максимально розрізняти взаємозв'язки між кліматичними умовами, агротехнічними показниками та остаточною врожайністю. Набір даних для цього дослідження зібраний з двох надійних джерел: 1) Продовольча та сільськогосподарська організація ООН (FAO), яка надає дані щодо врожайності різних культур та кількості використаних пестицидів; 2) Світовий банк, з бази якого отримані кліматичні дані, зокрема середньорічна температура та середня кількість опадів на рік для кожної країни.

Нами виконана візуалізація географічних даних включаючи показники врожайності для різних країн Європи, було створено карту, яка ілюструє розподіл культур та їх врожайність по країнах. На рис. 2.5 представлена карта Європи, що включає країни, для яких були доступні дані, зокрема детально відображена Україна. Це дає можливість проаналізувати просторовий розподіл даних та виявити можливості закономірності, наприклад, вплив кліматичних умов на врожайність певних культур у конкретних регіонах.

Щоб зрозуміти зв'язки між змінами, було виконано кореляційний аналіз між високими показниками, включаючи врожайність, кількість опадів, середню температуру та кількість використаних пестицидів. Нами написано код та створено гістограму, яка ілюструє топ-10 країн за кількістю записів у наборі даних (рис. 2.6). Встановлено, що найбільше записів наявно для Іспанії – понад 480 од.; для Італії – понад 320 од.; для Іспанії – понад 270 од.

Для побудови розподілу загалом використано 3591 записів щодо врожайності сільськогосподарських, що вказує на достатню вибірку для аналізу. Середній рівень врожайності становить 106640 гектограм з гектара. Це значення відображає загальну продуктивність культур у вибірці. Стандартне відхилення складає 111930 гектограм з гектара, що вказує на значне коливання врожайності в різних спостереженнях. Це свідчить про широкий діапазон продуктивності, що може бути зумовлено різними факторами, такими як тип культури, кліматичні умови, застосування агротехнічних заходів тощо.

На основі аналізу побудовано кореляційну матрицю (рис. 2.8), яка ілюструє взаємозв'язок між змінними. Наприклад, високі коефіцієнти кореляції між температурою та врожайністю певних культур можуть вказувати на сильний вплив температури на їх продуктивність. Кореляційна матриця містить коефіцієнти кореляції Пірсона, які показують ступінь зв'язку між парами змінних, з можливими значеннями від -1 до 1.

Для прогнозування врожайності сільськогосподарських культур було розроблено код для кількох моделей машинного навчання, включаючи лінійну регресію, модель випадкового лісу, XGBoost, ансамблеву модель та Gradient Boosting. Кожна з моделей має свої особливості, які допомагають забезпечити високу точність прогнозування на основі наявних даних.

Після розробки і навчання різних моделей прогнозування врожайності сільськогосподарських культур було проведено оцінку їхньої ефективності на тестових даних. Лінійна регресія показала середній рівень точності. Її простота робить цю модель швидкою та легкою для інтерпретації, але вона не завжди здатна врахувати складні, нелінійні зв'язки між факторами, що впливають на

врожайність. Випадковий ліс продемонстрував значно кращі результати порівняно з лінійною регресією. Завдяки ансамблю дерев рішень модель випадкового лісу може враховувати складні зв'язки та має високу стійкість до шуму в даних.

XGBoost показав найвищу точність серед усіх розглянутих моделей. Завдяки методам градієнтного підсилювання, ця модель досягла низького значення RMSE, що вказує на здатність ефективно враховувати як лінійні, так і нелінійні зв'язки в даних. Ансамблева модель об'єднала прогнози лінійної регресії, випадкового лісу та XGBoost. Завдяки цьому вона змогла об'єднати переваги всіх моделей, що призвело до хороших результатів. Gradient Boosting виявилася менш ефективною порівняно з XGBoost і випадковим лісом. Хоча Gradient Boosting здатна моделювати складні нелінійні залежності, вона вимагала більше обчислювальних ресурсів і часу на тренування.

Для визначення найефективнішої моделі для прогнозування врожайності сільськогосподарських культур було проведено порівняння п'яти моделей на основі метрик середньоквадратичної помилки (RMSE) та коефіцієнта детермінації (Accuracy). Найбільш ефективною моделлю за значенням RMSE є XGBoost із значенням 19657.25 та високим коефіцієнтом детермінації 0.9675. Це свідчить про те, що ця модель забезпечує найточніші прогнози серед розглянутих.

Нами розроблено заходи із охорони праці під час розроблення моделі машинного навчання для прогнозування врожайності сільськогосподарських культур, які забезпечать вимоги щодо охорони праці та безпеки у надзвичайних ситуаціях.

Результати визначення ефективності свідчать, що застосування моделі машинного навчання для прогнозування врожайності сільськогосподарських культур дозволяє отримати суттєвий економічний ефект у розмірі 160000 грн. за рахунок зменшення втрат врожаю та поточного планування агротехнічних заходів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ahmad, M.J.; Cho, G.H.; Kim, S.H.; Lee, S.; Adelodun, B.; Choi, K.S. Influence mechanism of climate change over crop growth and water demands for wheat-rice system of Punjab, Pakistan. *J. Water Clim. Chang.* 2021, 12, 1184–1202.
2. Ahmed, M.U.; Hussain, I. Prediction of Wheat Production Using Machine Learning Algorithms in northern areas of Pakistan. *Telecommun. Policy* 2022, 46, 102370.
3. Arshad, S.; Kazmi, J.H.; Javed, M.G.; Mohammed, S. Applicability of machine learning techniques in predicting wheat yield based on remote sensing and climate data in Pakistan, South Asia. *Eur. J. Agron.* 2023, 147, 126837.
4. Ashfaq, M.; Khan, I.; Alzahrani, A.; Tariq, M.U.; Khan, H.; Ghani, A. Accurate Wheat Yield Prediction Using Machine Learning and Climate-NDVI Data Fusion. *IEEE Access* 2024, 12, 40947–40961.
5. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32.
6. Fashoto, S.G.; Mbunge, E.; Ogunleye, G.; den Burg, J.V. Implementation of machine learning for predicting maize crop yields using multiple linear regression and backward elimination. *Precis. Agric.* 2021, 6, 679–697.
7. Fei, S.; Hassan, M.A.; Xiao, Y.; Su, X.; Chen, Z.; Cheng, Q.; Duan, F.; Chen, R.; Ma, Y. UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat. *Precis. Agric.* 2023, 24, 187–212.
8. Giraldo, P.; Benavente, E.; Manzano-Agugliaro, F.; Gimenez, E. Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy* 2019, 9, 352.
9. Haider, S.A.; Naqvi, S.R.; Akram, T.; Umar, G.A.; Shahzad, A.; Sial, M.R.; Khaliq, S.; Kamran, M. LSTM neural network based forecasting model for wheat production in Pakistan. *Agronomy* 2019, 9, 72.
10. Iqbal N., Shahzad M.U., Sherif E-S.M., Tariq M.U., Rashid J., Le T-V., Ghani A. Analysis of Wheat-Yield Prediction Using Machine Learning Models under



Climate Change Scenarios. Sustainability. 2024; 16(16):6976.  
<https://doi.org/10.3390/su16166976>

11. Ishaque, W.; Osman, R.; Hafiza, B.S.; Malghani, S.; Zhao, B.; Xu, M.; Ata-Ul-Karim, S.T. Quantifying the impacts of climate change on wheat phenology, yield, and evapotranspiration under irrigated and rainfed conditions. *Agric. Water Manag.* 2023, 275, 108017.

12. Kogan, F.; Guo, W.; Yang, W.; Shannon, H. Space-based vegetation health for wheat yield modeling and prediction in Australia. *J. Appl. Remote Sens.* 2018, 12, 026002.

13. Koval N., Tryhuba A., Kondysiuk I., Tryhuba I., Boiarchuk O., Rudynets M., Grabovets V., Onyshchuk V., Forecasting the Fund of Time for Performance of Works in Hybrid Projects Using Machine Training Technologies. Proceedings of the 3rd International Workshop on Modern Machine Learning Technologies and Data Science Workshop. Proc. 3rd International Workshop (MoMLLeT&DS 2021). Volume I: Main Conference. Lviv-Shatsk, Ukraine, June 5-6, 2021. pp.196-206.

14. Lee, C.C.; Zeng, M.; Luo, K. How does climate change affect food security? Evidence from China. *Environ. Impact Assess. Rev.* 2024, 104, 107324.

15. Li, Q.C.; Xu, S.W.; Zhuang, J.Y.; Liu, J.J.; Yi, Z.H.; Zhang, Z.X. Ensemble learning prediction of soybean yields in China based on meteorological data. *J. Integr. Agric.* 2023, 22, 1909–1927.

16. Lub, P., Tryhuba, A., Padyuka, R., Berezovetsky, S., Chubyk, R. Simulation modeling usage in the information system for the technological systems project management. *CEUR Workshop Proceedings*, 2023, 3453, pp. 139–148.

17. Paccioretti, P.; Bruno, C.; Gianinni Kurina, F.; Córdoba, M.; Bullock, D.; Balzarini, M. Statistical models of yield in on-farm precision experimentation. *Agron. J.* 2021, 113, 4916–4929.

18. Pang, A.; Chang, M.W.; Chen, Y. Evaluation of random forests (RF) for regional and local-scale wheat yield prediction in southeast Australia. *Sensors* 2022, 22, 717.

19. Rezaei, M.; Mousavi, S.R.; Rahmani, A.; Zeraatpisheh, M.; Rahmati, M.; Pakparvar, M.; Mahjenabadi, V.A.J.; Seuntjens, P.; Cornelis, W. Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. *Comput. Electron. Agric.* 2023, 209, 107821.
20. Rosenzweig, C.; Solecki, W.D.; Romero-Lankao, P.; Mehrotra, S.; Dhakal, S.; Ibrahim, S.A. *Climate Change and Cities: Second Assessment Report of the Urban Climate Change Research Network*; Cambridge University Press: Cambridge, UK, 2018.
21. Ruan, G.; Schmidhalter, U.; Yuan, F.; Cammarano, D.; Liu, X.; Tian, Y.; Zhu, Y.; Cao, W.; Cao, Q. Exploring the transferability of wheat nitrogen status estimation with multisource data and Evolutionary Algorithm-Deep Learning (EA-DL) framework. *Eur. J. Agron.* 2023, 143, 126727.
22. Sahbeni, G.; Székely, B.; Musyimi, P.K.; Timár, G.; Sahajpal, R. Crop Yield Estimation Using Sentinel-3 SLSTR, Soil Data, and Topographic Features Combined with Machine Learning Modeling: A Case Study of Nepal. *AgriEngineering* 2023, 5, 1766–1788.
23. Sarkodie S.A., Adams S. Electricity access, human development index, governance and income inequality in Sub-Saharan Africa. *Energy Rep*, 6 (2020), pp. 455-466.
24. Silva, A.A.D. Investigation into the Effects of Microbial Communities on Biogeochemical Cycles in Soil. Ph.D. Thesis, Université Catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium, 2013.
25. Sinwar, D.; Dhaka, V.S.; Sharma, M.K.; Rani, G. AI-based yield prediction and smart irrigation. In *Internet of Things and Analytics for Agriculture*; Springer: Singapore, 2020; Volume 2, pp. 155–180.
26. Srivastava, R.K.; Mequanint, F.; Chakraborty, A.; Panda, R.K.; Halder, D. Augmentation of maize yield by strategic adaptation to cope with climate change for a future period in Eastern India. *J. Clean. Prod.* 2022, 339, 130599.

27. Tariq, A.; Yan, J.; Gagnon, A.S.; Riaz Khan, M.; Mumtaz, F. Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. *Geo-Spat. Inf. Sci.* 2023, 26, 302–320.
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., Tryhuba, I., Malanchuk, O., Marmulyak, A. A deep neural network model for predicting the competitive score of social projects for community development. *CEUR Workshop Proceedings*, 2024, 3711, pp. 55–74.
30. Tryhuba, A., Malanchuk, O., Tryhuba, I. Prediction of the Duration of Inpatient Treatment of Diabetes in Children Based on Neural Networks. *CEUR Workshop Proceedings*, 2023, 3426, pp. 122–135.
31. Tryhuba, I., Tryhuba, A., Grabovets, V., Bodak, V., Horodetska1, N. Forecasting the duration of work in plant protection projects. *CEUR Workshop Proceedings*, 2023, 3453, pp. 96–105.
32. Tryhuba, I., Tryhuba, A., Hutsol, T., Tulej, W., Sojak, M. Prediction of Biogas Production Volumes from Household Organic Waste Based on Machine Learning. *Energies*, 2024, 17(7), 1786.
33. Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* 2020, 177, 105709.
34. Wadoux, A.M.C.; Brus, D.J.; Heuvelink, G.B. Sampling design optimization for soil mapping with random forest. *Geoderma* 2019, 355, 113913.
35. Zhang, L.; Traore, S.; Ge, J.; Li, Y.; Wang, S.; Zhu, G.; Cui, Y.; Fipps, G. Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel. *Comput. Electron. Agric.* 2019, 166, 105031.
36. Zhang, M.; Gao, Y.; Zhang, Y.; Fischer, T.; Zhao, Z.; Zhou, X.; Wang, Z.; Wang, E. The contribution of spike photosynthesis to wheat yield needs to be considered in process-based crop models. *Field Crop. Res.* 2020, 257, 107931.

37. Введення в машинне навчання за допомогою Python и Scikit-Learn. URL: <https://habr.com/ua/company/mlclass/blog/247751/> (дата звернення: 10.05.2024).
38. Жидецький В.Ц., Джигирей В.С., Мельников О.В. Основи охорони праці. Підручник. Вид. 5-е, доповнене. Львів: Афіша, 2012. 350с.
39. Лехман С.Д., Рублев В.І., Рябцев Б.І. Запобігання аварійності і травматизму у сільському господарстві. К.: Урожай, 1993. 267 с.
40. Плєскач В.Л., Рогушина Ю.В., Кустова Н.П. Інформаційні технології та системи. К.: Книга, 2004. 519 с.
41. Tryhuba A., Ivanyshyn V., Chaban V., Mushenyk I., Zharikova O. Computer model of resource demand planning for dairy farms. Independent Journal of Management & Production (Special Edition ISE, S&P). 2021. 12(3), pp. 138-149. URL: <http://www.ijmp.jor.br/index.php/ijmp/article/view/1531/1971> (Last accessed: 16.10.2024).
42. 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 (MoMLeT+DS 2020). 2020. I. P. 327-340.