

ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
(повне найменування вищого навчального закладу)
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ
(повне найменування інституту, назва факультету (відділення))
КАФЕДРА ПРОГРАМНИХ ЗАСОБІВ ТА ТЕХНОЛОГІЙ
(повна назва кафедри (предметної, циклової комісії))

Пояснювальна записка

до кваліфікаційної роботи
бакалавра
(освітній рівень)

на тему: «Розробка програмного забезпечення для прогнозування хвороб
рослин»

Виконав: здобувачка освіти 4 року навчання,
групи 4ПР4 спеціальності
121 - «Інженерія програмного забезпечення»
(шифр і назва спеціальності)

Болдирєва Таїсія Сергіївна
(прізвище та ініціали)

Керівник д.т.н., проф. Шерстюк В.Г.
(прізвище та ініціали)

Рецензент к.т.н. доцент Вишемирська С.В.
(прізвище та ініціали)

Хмельницький - 2024

Херсонський національний технічний університет

(повне найменування вищого навчального закладу)

Факультет, відділення Інформаційних технологій та дизайну
 Кафедра Програмних засобів і технологій
 Освітній рівень бакалавр
 Спеціальність 121 – Інженерія програмного забезпечення
 (шифр і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри
Програмних засобів та технологій
к.т.н. доц. О.Є. Огнєва
 “ ____ ” _____ 2024 р.

З А В Д А Н Н Я**НА КВАЛІФІКАЦІЙНУ РОБОТУ ЗДОБУВАЧУ**Болдирєвій Таїсії Сергіївні

(прізвище, ім'я, по батькові)

Тема роботи «Розробка програмного забезпечення для прогнозування хвороб рослин»

керівник роботи д.т.н. професор Шерстюк В.Г.,
 (прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджена наказом вищого навчального закладу від 09.02.2024 р. №220-С

2. Строк подання здобувачем роботи 17.06.2024 р.

3. Вихідні дані до роботи літературні та періодичні джерела,
дані з метеостанції METOS by Pessl Instruments із використанням IoT
платформи FieldClimate

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

1) сучасний стан науково-прикладних розробок у галузі прогнозування вірогідності хвороб сільськогосподарських культур;

2) розробка методу та алгоритму вирішення задачі;

3) результати використання алгоритмів машинного навчання під час прогнозування вірогідності виникнення хвороб сільськогосподарських культур.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

1. архітектура системи прогнозування.

2. матриця кореляції.

3. моделі прогнозування.

4. алгоритм отримання оптимальної моделі прогнозу захворювання.

5. порівняння досягнутих результатів використання моделей.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів виконання роботи	Термін виконання етапів роботи	Примітки
1.	Отримання завдання	09.02.2024	Виконано
2.	Підбір літератури	09.02.2024-21.02.2024	Виконано
3.	Аналіз предметної області	22.02.2024-28.02.2024	Виконано
4.	Розробка та обґрунтування завдання	01.03.2024-14.03.2024	Виконано
5.	Розробка концептуальної моделі	15.03.2024-21.03.2024	Виконано
6.	Розробка алгоритму	22.03.2024-04.04.2024	Виконано
7.	Проектування програмної системи	05.04.2024-11.04.2024	Виконано
8.	Розробка інтерфейсу програмної системи	12.04.2024-18.03.2024	Виконано
9.	Тестування програмної системи	19.04.2024-25.04.2024	Виконано
10.	Оформлення пояснювальної записки	26.04.2024-09.05.2024	Виконано
11.	Захист кваліфікаційної роботи	17.06.2024	Виконано

Здобувачка вищої освіти _____ Болдирєва Т.С.
(підпис) (прізвище та ініціали)Керівник роботи _____ Шерстюк В.Г.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Кваліфікаційна робота бакалавра: 81 сторінка, 32 рисунка, 6 таблиць, 1 додаток, 18 джерел.

Мета роботи – визначення найбільш ефективного методу відповідно до метрик RMSE та R^2 , який буде інтегровано в програмний застосунок для надання прогнозів ймовірності захворювання спостережуваних сільськогосподарських культур в режимі реального часу.

Об’єкт дослідження – процеси збору та обробки даних щодо прогнозування вірогідності виникнення хвороби сільськогосподарських культур.

Предмет дослідження – алгоритми машинного навчання для прогнозування вірогідності виникнення хвороби певних сортів кукурудзи.

Методи дослідження – методи лінійної регресії, нейронної мережі прямого поширення, Random Forest, критичного аналізу і логічного узагальнення у галузі машинного навчання, методи статистичного аналізу даних.

Новизна роботи полягає у встановленні набору вхідних вимірюваних кліматичних даних та оцінці ефективності алгоритмів машинного навчання під час прогнозування вірогідності виникнення хвороб кукурудзи, що дає змогу оптимізувати вибір підходу під час проектування конкретних типів систем із обліком метрик R^2 та RMSE.

Практична цінність результатів роботи полягає в розробці програмного застосунку на базі розробленого підходу, який допоможе аграріям прогнозувати виникнення певних хвороб кукурудзи, що дозволить своєчасно вжити необхідних заходів для захисту врожаю.

Ключові слова: *сільськогосподарські культури, захворювання, прогнозування, машинне навчання, лінійна регресія, нейронна мережа, Random Forest.*

АНОТАЦІЯ

Кваліфікаційна робота бакалавра складається зі вступу, трьох розділів, висновку, списку використаних джерел та додатку.

У роботі досліджено алгоритми машинного навчання для прогнозування вірогідності виникнення хвороби певних сортів кукурудзи.

В результаті виконання роботи розроблено програмну реалізацію алгоритму машинного навчання для прогнозування ймовірності виникнення хвороби кукурудзи Fusarium Head Blight.

Результати роботи подано у вигляді, що дозволяє побачити та оцінити безпосереднє використання моделі машинного навчання.

Розроблене програмне рішення може бути інтегровано в програмний застосунок для надання прогнозів щодо ймовірності захворювання спостережуваних культур в режимі реального часу.

Оцінка ефективності запропонованого рішення може бути поширена на інші зернові культури та інші хвороби.

Результати проведених досліджень сприятимуть зменшенню втрат врожаю, підвищенню продовольчої безпеки та експортному потенціалу.

ABSTRACT

The bachelor's thesis consists of an introduction, three chapters, a conclusion, a list of sources and an appendix.

The thesis examines machine learning algorithms for predicting the probability of disease occurrence of certain varieties of corn.

As a result of the thesis, a software implementation of the machine learning algorithm was developed to predict the probability of occurrence of Fusarium Head Blight corn disease.

The results of the thesis are presented in a form allowing to see and evaluate the direct use of the machine learning model.

The developed software solution can be integrated into a software application to provide predictions about the probability of disease of observed crops in real time.

Evaluation of the effectiveness of the proposed solution can be extended to other grain crops and other diseases.

The results of the conducted research will contribute to the reduction of crop losses, increase in food security and export potential.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	9
ВСТУП	10
РОЗДІЛ 1. СУЧАСНИЙ СТАН НАУКОВО-ПРИКЛАДНИХ РОЗРОБОК У ГАЛУЗІ ПРОГНОЗУВАННЯ ВІРОГІДНОСТІ ХВОРОБ С/Г КУЛЬТУР	13
1.1. Аналіз вимог щодо комп'ютеризованого моніторингу ґрунто-кліматичних параметрів.....	13
1.2. Аналіз відомих ML алгоритмів.....	16
1.3. Аналіз сучасних програмно-апаратних рішень прогнозування вірогідності виникнення хвороб с/г культур.....	25
1.4. Обґрунтування мети та задач дослідження	30
1.5. Висновки.....	31
РОЗДІЛ 2. РОЗРОБКА МЕТОДУ ТА АЛГОРИТМУ ВИРІШЕННЯ ЗАДАЧІ	33
2.1. Опис обмежень досліджуваної технології.....	33
2.2. Узагальнена структурно-алгоритмічна організація досліджуваної технології	34
2.3. Методика проведення досліджень.....	36
2.4. Висновки.....	51
РОЗДІЛ 3. РЕЗУЛЬТАТИ ВИКОРИСТАННЯ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ ПІД ЧАС ПРОГНОЗУВАННЯ ВІРОГІДНОСТІ ВИНИКНЕННЯ ХВОРОБ С/Г КУЛЬТУР.....	52
3.1. Регресійні метрики для оцінки якості моделей ML	52
3.2. Результати для методу лінійної регресії	53
3.3. Результати для алгоритму Random Forest.....	55
3.4. Результати використання нейронної мережі прямого поширення.....	59
3.5. Порівняльний аналіз отриманих результатів	63

3.6. Перспективи подальшого вдосконалення.....	66
3.7. Висновки.....	68
ВИСНОВКИ.....	70
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	72
ДОДАТОК А. ПРОГРАМНИЙ КОД.....	74

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

CNN – convolutional neural network

EMA – exponential moving average

FNN – feedforward neural network

HTTP – hypertext transfer protocol

IoT – internet of things

KNN – k-nearest neighbors

LSTM – long short-term memory

MIoT – massive internet of things

ML – machine learning

RNN – recurrent neural network

SVM – support vector machines

WSN – wireless sensor networks

НМ – нейронна мережа

ШІ – штучний інтелект

ВСТУП

Масивний Інтернет речей (MІoT) – це мережа, що характеризується великою кількістю взаємопов’язаних пристроїв. Як правило, ці пристрої, часто розташовані у віддалених місцях, збирають дані і передають інформацію на центральний сервер або хмару. Ці «речі», як правило, недорогі, малопотужні і мають обмежені обчислювальні можливості. Наочним прикладом є сільськогосподарські датчики, розміщені на полях, річках і в лісах.

Стрімке зростання MІoT значною мірою пов’язане з розвитком стільникового зв’язку. За прогнозами Juniper, кількість підключень до IoT зросте з 35 мільярдів у 2020 році до 83 мільярдів в 2024 році, причому приблизно половина з них припадатиме на категорію MІoT. Водночас IndustryARC прогнозує середньорічний темп зростання (CAGR) ринку масового IoT на рівні 7,1%, який досягне 121,4 мільярда доларів до 2026 року.

Зростання світового попиту на продовольство, який, за прогнозами, подвоїться до 2050 року, ще більше загострює проблему підвищення продуктивності сільського господарства.

Незважаючи на такі виклики, як зниження рівня води, зміна клімату та зменшення площ орних земель, технології, що базуються на даних, є перспективним шляхом підвищення продуктивності фермерських господарств на 67% до 2050 року, згідно з даними Міжнародного науково-дослідного інституту продовольчої політики (International Food Policy Research Institute).

Польові випробування продемонстрували ефективність сенсорних технологій, таких як точне зрошення, яке може підвищити продуктивність фермерських господарств на 45%, зменшивши при цьому споживання води на 35%. Аналогічно, використання сенсорних даних з систем аерофотозйомки, таких як дрони, дозволяє фермерам картографувати поля, стежити за станом посівів і виявляти аномалії. Таким чином, потенціал MІoT знаходить гідне

застосування в аграрному секторі, де потреба в освоєнні великих і віддалених територій є нагальною.

Однак, для того щоб з часом накопичені дані полегшили визначення оптимальних методів ведення сільського господарства, що призводить до підвищення врожайності, зменшення витрат ресурсів та мінімізації впливу на навколишнє середовище, потрібно їх певним чином обробляти.

Всупереч традиційним уявленням, ШІ відіграє ключову роль у сучасному сільському господарстві. Тонкощі фермерства, з його переплетеною павутиною передсезонних і сезонних рішень, виграють від здатності ШІ обробляти великі обсяги даних від сільськогосподарського обладнання, датчиків навколишнього середовища і віддалених джерел. Така зміна парадигми підвищує ефективність і стійкість у всьому ланцюжку виробництва продуктів харчування.

Підвищення врожайності та зменшення витрат ресурсів набуває особливої нагальності в контексті зменшення площ вирощування сільськогосподарських культур в Україні. Тому значно актуалізується проблема розробки та впровадження стратегій менеджменту сільськогосподарських підприємств на основі логіки «більші обсяги виробництва на менших площах». Такий результат може бути досягнутий шляхом збільшення показників збереження сільськогосподарських культур протягом повного циклу вирощування, шляхом прогнозування появи хвороб сільськогосподарських культур.

Отже, тема даної кваліфікаційної роботи є актуальною, оскільки в рамках дослідження визначається найбільш ефективний ML алгоритм для прогнозування вірогідності виникнення хвороби кукурудзи.

Метою роботи є визначення найбільш ефективного методу відповідно до метрик RMSE та R^2 , який буде інтегровано в програмний застосунок для надання прогнозів ймовірності захворювання спостережуваних сільськогосподарських культур в режимі реального часу.

Об'єктом дослідження є процес збору та обробки даних щодо прогнозування вірогідності виникнення хвороби сільськогосподарських культур.

Предметом дослідження є алгоритм машинного навчання для прогнозування

вірогідності виникнення хвороби певних сортів кукурудзи.

Застосовано наступні методи дослідження: методи лінійної регресії, нейронної мережі прямого поширення, Random Forest, критичного аналізу і логічного узагальнення у галузі машинного навчання, методи статистичного аналізу даних.

Новизна роботи полягає у встановленні набору вхідних вимірюваних кліматичних даних та оцінці ефективності алгоритмів машинного навчання під час прогнозування вірогідності виникнення хвороб кукурудзи, що дає змогу оптимізувати вибір підходу під час проектування конкретних типів систем із обліком метрик R^2 та RMSE.

Практична цінність результатів роботи полягає в розробці програмного застосунку на базі розробленого підходу, який допоможе аграріям прогнозувати виникнення певних хвороб кукурудзи, що дозволить своєчасно вжити необхідних заходів для захисту врожаю.

Структура роботи: пояснювальна записка: 78 стор., 32 рис., 6 таблиць, 1 додаток, 18 джерел.