Industry Oriented Mini Project Report

on

INTEGRATED AGRICULTURAL DECISION SUPPORT SYSTEM (IADSS)

Submitted in partial fulfillment of the requirements
for the award of degree of

BACHELOR OF TECHNOLOGY

in

Information Technology

by

Yerrolla Jansi (20WH1A1208)

Murikipudi Sannitya (20WH1A1213)

Koravat Padma (20WH1A1216)

Under the esteemed guidance of

Mr. K. Srikar Goud

Assistant Professor

Department of Information Technology

BVRIT HYDERABAD College of Engineering for Women

Rajiv Gandhi Nagar,Nizampet Road, Bachupally,Hyd–500090

(Affiliated to Jawaharlal Nehru Technological University, Hyderabad)

(NAAC 'A' Grade & NBA Accredited- ECE, EEE, CSE & IT)

December 12, 2023

BVRIT HYDERABAD
College of Engineering for Women
Rajiv Gandhi Nagar, Nizampet Road, Bachupally, Hyderabad – 500090
(Affiliated to Jawaharlal Nehru Technological University Hyderabad)
(NAAC 'A' Grade & NBA Accredited- ECE, EEE, CSE & IT)

CERTIFICATE

This is to certify that the Project report on **"Integrated Agricultural Decision Support System (IADSS)"** is a bonafide work carried out by **Yerrolla Jansi (20WH1A1208)** , **Murikipudi Sannitya (20WH1A1213)** , **Koravat Padma (20WH1A1216)** in the partial fulfillment for the award of B.Tech degree in **Information Technology , BVRIT HYDERABAD College of Engineering for Women, Bachupally, Hyderabad** affiliated to Jawaharlal Nehru Technological University, Hyderabad, under my guidance and supervision. The results embodied in the project work have not been submitted to any other university or institute for the award of any degree or diploma.

Internal Guide	**Head of the Department**
Mr. K. Srikar Goud	**Dr. Aruna Rao S L**
Assistant Professor	**Professor & HoD**
Department of IT	**Department of IT**

External Examiner

DECLARATION

We hereby declare that the work presented in this project entitled **"Integrated Agricultural Decision Support System (IADSS)"** submitted towards completion of in IV year I sem of B.Tech IT at "BVRIT HYDERABAD College of Engineering for Women", Hyderabd is an authentic record of our original work carried out under the esteemed guidance of **Mr. K. Srikar Goud, Assistant Professor** Department of Information Technology.

Yerrolla Jansi (20WH1A1208)

Murikipudi Sannitya (20WH1A1213)

Koravat Padma(20WH1A1216)

This project report is dedicated to my beloved Family members and supervisor for their limitless support and encouragement and to you as a reader

ACKNOWLEDGMENTS

We would like to express our profound gratitude and thanks to **Dr. K. V. N. Sunitha, Principal, BVRIT HYDERABAD College of Engineering for Women** for providing the working facilities in the college.

Our sincere thanks and gratitude to **Dr. Aruna Rao S L, Professor & Head, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for all the timely support, constant guidance and valuable suggestions during the period of our project.

We are extremely thankful and indebted to our internal guide, **Mr. K. Srikar Goud, Assistant Professor, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for his constant guidance, encouragement and moral support throughout the project.

Finally, we would also like to thank our Project Coordinators **Ms. K.S. Niraja, Assistant Professor and Ms. A. Aruna Jyothi, Assistant Professor**, all the faculty and staff of Department of IT who helped us directly or indirectly, parents and friends for their cooperation in completing the project work.

Yerrolla Jansi(20WH1A1208)

Murikipudi Sannitya(20WH1A1213)

Koravat Padma(20WH1A1216)

ABSTRACT

Agriculture, serving as the backbone of the Indian economy, stands as a cornerstone in the livelihoods of millions, contributing significantly to the economy, providing jobs, and meeting the country's food needs.Despite its vital role, the sector faces many challenges like inadequate access to real-time weather forecasts, inefficient water management, imprecise fertilization practices, and a lack of data-driven insights for crop management that hinder its growth and efficiency. "Integrated Agricultural Decision Support System (IADSS)" integrates agriculture and machine learning, providing a comprehensive platform for farmers. Utilizing advanced meteorological data and cutting-edge algorithms, it delivers precise, location-specific weather forecasts, enabling informed decisions on planting, irrigating, and harvesting. IADSS goes beyond weather forecasting, offering crop selection guidance based on historical data and predicting optimal fertilizers through machine learning models. The system's proactive features empower farmers to take preventive measures, reducing reliance on chemical interventions for sustainable agriculture. Accessible via smartphones, tablets, and computers, the user-friendly web application equips farmers with data-driven insights for crop cultivation and fertilizer management, ultimately enhancing global agricultural productivity, sustainability, and resilience in the face of a changing climate. The project holds promise in improving the livelihoods of farmers worldwide by providing essential tools for success in a dynamic agricultural landscape.

Keywords:crop prediction, fertilizer recommendation, weather forecast, machine learning.

Contents

Declaration **i**

Acknowledgement **iii**

Abstract **iv**

List of Figures **vii**

List of Tables **viii**

1 Introduction **1**
1.1 Motivation 2
1.2 Objective 3
1.3 Problem Definition 3

2 Literature Survey **4**

3 System Design **8**
3.1 UML Diagrams 9
3.1.1 Use Case Diagram 9
3.1.2 Sequence Diagram 10
3.2 Tools and Libraries used 11

4 Methodology **12**
4.1 Proposed System 12
4.2 Dataset 13
4.3 Data Pre-processing 15
4.4 Training and Testing 17
4.4.1 Random Forest Classifier 17
4.4.2 Decision Tree Classifier 18
4.4.3 Support vector machine 18
4.4.4 Naive Bayes Classifer 19

4.5 Prediction 21

5 Implemention **22**

6 Results and Discussions **25**

7 Conclusion and Future Scope **27**

List of Figures

3.1 Use Case Diagram 10
3.2 Sequence Diagram 11

4.1 Architecture 13
4.2 Attributes of the Dataset 14
4.3 Encoding of attributes 16
4.4 Inverse Encoding of attributes 16
4.5 Structure of Random Forest 17
4.6 Model Defination of DecisionTree Classifier 18
4.7 Structure of Support vector algorithm 19
4.8 Structure of Naive Bayes Algorithm 19
4.9 importing All Algorithms classifiers 20
4.10 Training and Saving the model 20
4.11 Output for given input values in trained model 21

5.1 Web Output taking input values 22
5.2 Output Representing Wheat Crop 23
5.3 Output Representing NPK Fertilizer 23
5.4 Recommended crop and geographical location weather forecast . 24
5.5 Recommended fertilizer for crop predicted 24

6.1 Comparision chart of all algorithms based on metrics 25
6.2 Comparison of all algorithms based on metrics 26

List of Tables

4.1 Features and their description 15
6.1 Metrics Values for Different Algorithms 26

Chapter 1

Introduction

The Integrated Agricultural Decision Support System (IADSS) represents a groundbreaking innovation poised to transform the agricultural landscape. It serves as a convergence point, seamlessly integrating cutting-edge technological advancements with the intricate demands of modern agriculture.

At its core, IADSS consolidates three vital components -weather forecasting, crop prediction, and personalized fertilizer recommendations into a singular, cohesive platform. This amalgamation harnesses the power of data-driven insights, leveraging sophisticated algorithms and advanced data sources to provide farmers with actionable information crucial for optimizing their agricultural practices.

By bridging the gap between technology and agriculture, IADSS introduces a paradigm shift in how farmers approach decision-making processes. It transcends traditional methods by offering a comprehensive solution that addresses multiple facets of agricultural challenges, empowering farmers with the tools and knowledge necessary to navigate the complexities of contemporary farming.

This innovative system doesn't merely provide standalone solutions but rather orchestrates a synergy among various agricultural elements. It integrates weather forecasts tailored to specific locations, predictive analytics for crop suitability based on historical data, and personalized fertilizer recommendations crafted through machine learning models, all within a single accessible platform.

The significance of IADSS extends beyond its technological prowess. It represents a commitment to enhancing agricultural productivity, sustainability, and resilience. By empowering farmers with data-driven insights, IADSS aims to not only optimize crop yields but also minimize risks associated with adverse weather conditions, reduce reliance on chemical interventions, and contribute to the cultivation of a more sustainable and resilient agricultural ecosystem.

In essence, IADSS embodies the fusion of technological innovation and agricultural expertise, paving the way for a new era in farming practices, where informed decisions based on precise data are the cornerstone of success.

1.1 Motivation

The motivation behind the development of the Integrated Agricultural Decision Support System (IADSS) stems from the urgent need to address the multifaceted challenges that modern farmers encounter within an ever-evolving agricultural environment.

Climate Uncertainties: Contemporary agriculture operates within a climate landscape that's becoming increasingly unpredictable. Unforeseen weather fluctuations, extreme events, and shifting patterns pose significant risks to crop cultivation. These uncertainties threaten crop yields, affecting not only the farmer's income but also global food security.

Risk Mitigation: IADSS emerges as a response to these challenges, recognizing the critical role technology can play in mitigating risks. By providing farmers with precise, real-time weather forecasts tailored to their specific locations, the system empowers them to make informed decisions. This includes optimal times for planting, irrigation, and harvesting, mitigating potential losses due to adverse weather conditions.

Optimized Crop Choices: Another motivation for IADSS lies in the capability to guide farmers in selecting crops best suited to their unique environmental conditions. By leveraging historical data and predictive analytics, the system assists in making informed choices about which crops will thrive in a particular region. This empowers farmers to optimize their agricultural productivity and adapt to changing environmental conditions.

Sustainability Promotion: Moreover, IADSS promotes sustainable agricultural practices. By offering personalized fertilizer recommendations based on soil tests and crop preferences, the system enables farmers to adopt preventive measures. This reduces the need for chemical interventions, minimizing environmental impact while fostering sustainable farming practices.

Empowerment Through Accessibility: Crucially, IADSS is designed to be accessible. It aims to bridge the technological gap by providing a user-friendly interface accessible via smartphones, tablets, and computers. This ensures that farmers, irrespective of their technological proficiency, can benefit from the valuable insights and recommendations offered by the system.

In essence, the motivation behind IADSS is rooted in empowering farmers with the necessary tools to navigate the uncertainties of modern agriculture. It seeks to minimize risks, optimize productivity, and promote sustainable farming practices, ultimately contributing to the resilience and prosperity of agricultural communities worldwide.

1.2 Objective

The primary objective of the Integrated Agricultural Decision Support System (IADSS) is to revolutionize the agricultural sector by seamlessly integrating cutting-edge technological advancements with the intricate demands of modern agriculture. At its core, IADSS aims to consolidate weather forecasting, crop prediction, and personalized fertilizer recommendations into a unified, cohesive platform. By leveraging data-driven insights, sophisticated algorithms, and advanced data sources, the system seeks to provide actionable information crucial for optimizing agricultural practices.

1.3 Problem Definition

IADSS is designed to tackle the complex challenges facing modern agriculture, particularly the escalating uncertainties in climate patterns that pose a threat to global food security. The system's core mission is to empower farmers by delivering accurate weather forecasts, enabling informed decisions on planting, irrigation, and harvesting. It addresses critical issues such as crop selection, sustainability, and accessibility, guiding farmers in choosing the most suitable crops, advocating for sustainable practices through personalized fertilizer recommendations, and ensuring user-friendly access. In essence, IADSS strives to mitigate climate-related risks, optimize agricultural practices, and enhance the resilience and prosperity of farming communities on a global scale.

Chapter 2

Literature Survey

The literature survey explores diverse applications of machine learning and data-driven approaches in agriculture, aiming to revolutionize traditional farming practices. Leveraging algorithms like Decision Tree, Naive Bayes, and Random Forest, the research focuses on predicting crop yields, offering fertilizer recommendations, and aiding crop selection decisions. The integration of advanced technologies, including sensors, IoT, and mobile applications, underscores a commitment to empower farmers, enhance productivity, and address challenges in the agricultural landscape. This comprehensive study seeks to contribute to the prosperity of farmers, optimize agricultural outcomes, and underscore the pivotal role of data science in shaping the future of agriculture

The paper[1] predicts crops to be sown based on soil where the input is taken from sensors and based on dataset and values using decision tree supervised machine learning algorithms such as Naive Bayes theorem,Random Forest,Decision Trees etc where prediction of crop yield includes forecasting factors like temperature, humidity, rainfall, etc and crop yield based on soil moisture includes few measures like NPK (Nitrogen, Phosphorous and potassium) and pH values using various sensors.

The paper[2] proposes a user-friendly yield prediction system for Indian farmers, addressing the challenge of low crop yields compared to international standards, which contributes to issues like farmer suicides. The system, accessible through a mobile application, utilizes GPS for user location and allows input of area and soil type. Machine learning algorithms, including Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN), predict crop yield or suggest the most profitable crops. Random Forest demonstrated the highest accuracy at 95%. The system also advises on optimal fertilizer application timing, aiming to enhance crop yield. Future work includes dataset updates for accurate predictions, process automation, and integrating fertilizer recommendations based on soil and climate analysis.

The paper[3] proposes a machine learning solution, specifically the Naive Bayes Gaussian classifier with boosting algorithm, to assist beginner farmers in predicting and optimizing crop yields amid unpredictable climatic changes. The system collects seed data with key parameters and develops an Android application for easy input. The research aims to provide accurate crop predictions, helping farmers with limited knowledge. Future enhancements include fertilizer suggestions, cropland guidelines, and crop health monitoring, contributing to a sustainable agricultural future. In recent years the erratic weather changes have led to various polemics. Indonesia as an agricultural country with a large part of the population earning a living in agriculture feels real impacts such as crop failure. The use of cropping patterns that have been carried out from generation to generation without regard to climate change and the environment is the cause. Technology changes in agriculture needs by utilizing government-owned data to help farmers by providing recommendations for food crops. This study aims to use a data classification technique with the Naïve Bayes algorithm in obtaining the results of the recommendation of food plant types. Data obtained from the Provincial Office of DI. Yogyakarta. The parameters considered are the weather, yields, and selling prices of the four districts in the province. Data selection and data cleaning are needed to retrieve attributes that affect the results of recommendations. To find out the performance of Naïve Bayes to the dataset, use WEKA. The cross-validation method is used to validate data. The results showed an accuracy of 85.71%. Naïve Bayes is feasible to be used for a dataset of recommended crop species, supported by the results of the sensitivity 0.857 and a specificity of 0.862 as validation. Naïve Bayes consistency is consistent with the Kappa Statistics value of 0.8084. Besides, the method error when classifying and the time has taken also calculated.

The paper[4] proposes The crop recommendation system employing machine learning methods,based on several factors, including nitrogen (N), phosphorus (P), potassium (K) and humidity, we will advise the best crop for the given site. We analyzed various algorithms like KNN, Decision Tree, Random Forest, SVM etc. But based on various accuracy levels we committed to random forest implementation. Means, In this paper they used random forest algorithm which allowed to train upon large dataset and the performance of the recommendation system is measured using accuracy score.

The paper[5]proposes progressions in machine learning and crop simulation techniques. Machine learning (ML), on the other hand, aims to make forecast by discovering associations between input and response variables. Various elements, including weather and soil, are making it challenging for farmers to cultivate crops. Developing effective agricultural and food policies on a regional and international scale requires accurate crop yield forecasts. Our proposed solution combines two machine learning algorithms to optimize agriculture by predicting crop yield and recommending fertilizer. This script is innovative because it allows the user to predict the most suitable crop based on basic information such as soil characteristics and weather conditions. We have utilized Random Forest and Logistic Regression for the system's implementation. This model serves as an example of hybrid ML approaches which could solve the above mentioned issues and increase the yield.

The paper [6] proposes integration of advanced technologies such as advanced sensors coupled with Internet of Things (IoT) could escalate the agricultural production and minimize the economic loss. Studies have been conducted across the world that satisfactorily demonstrated the implication of integrated IoT-smart sensors in monitoring environmental factors such as moisture, humidity, temperature, and soil composition that are critical for crop growth. Green house gases such as Carbon dioxide, Methane, etc., are also measured through automated sensors. Smart farming also enables measurement of nitrogen contents in soil that helps farmers to determine the amount of fertilizers to be used in farm lands.

The paper [7] proposes a decision tree supervised machine learning model to address key issues in agriculture planning, specifically focusing on soil moisture for crop yield prediction. Recognizing the economic impact of modern farming, the research highlights the importance of factors like soil nutrients, crop prediction, and monitoring. The model considers parameters such as temperature, humidity, and pH values, using various sensors for accurate predictions. The proposed system aids farmers in deciding the type of crop based on soil moisture values, contributing to economic growth and maximizing crop yield. The extension of this work includes automatic detection of crop yield by incorporating additional parameters like weather forecast and soil testing for improved agricultural outcomes.

This paper [8] introduces a comprehensive approach to addressing the economic significance of agriculture in India by proposing an advanced system for crop yield and fertilizer prediction. Utilizing data mining and machine learning, the study analyzes past weather and technological factors to accurately forecast future crop yield. The Random Forest regression model is employed for precise crop yield predictions, while the Decision Tree algorithm forecasts fertilizer requirements. The user interface, developed with Tkinter, allows users to input critical factors like state, district, year, season, crop, and area, enhancing accessibility for accurate predictions. Notably, the Random Forest Regression model demonstrates superior accuracy compared to other algorithms. The proposed system aims to empower agronomists, farmers, and policymakers, offering insights to maximize crop yield and optimize fertilizer use, thereby contributing to overall economic growth. The user-friendly interface promotes accessibility, and future enhancements, including the development of a mobile application, are suggested for more effective utilization.

The paper [9] introduces a comprehensive strategy for addressing the pivotal role of agriculture in India's socio-economic landscape, with a focus on optimizing crop productivity. Understanding the challenges faced by farmers in predicting yields amidst environmental fluctuations, the study employs deep learning algorithms like CNN, alongside machine learning models such as SVM, Naive Bayes, Random Forest, and XG Boost. Utilizing two datasets for fertiliser prediction and crop recommendation, the paper advocates for an ensemble technique to enhance prediction accuracy. The study's primary goal is to empower farmers by providing insights into crop production, disease prediction, fertiliser recommendations, and optimal crop choices. Notably, the paper underscores the accuracy of algorithms like Random Forest and CNN, comparing favorably with previous techniques. Additionally, it outlines the development of front-end and back-end frameworks using React JS and Django, facilitating user-friendly applications for farmers. The ultimate aim is to contribute to the prosperity of both farmers and the nation's agricultural output, with future plans to integrate cutting-edge features and expand datasets with new qualities.

This paper [10] underscores the significance of agriculture in India's economy and the challenges posed by natural calamities, often leading to financial losses for farmers and unfortunate outcomes like suicides. To address this, the study proposes an intelligent crop recommendation system, leveraging machine learning algorithms including Decision Tree, Naïve Bayes, Support Vector Machine, Logistic Regression, Random Forest, and XGBoost. The system aims to guide farmers in making informed decisions about crop selection based on essential parameters like Nitrogen, Phosphorous, Potassium, pH Value, Humidity, Temperature, and Rainfall. The research envisions increased productivity, profitability, and overall economic well-being, emphasizing the crucial role of data science in optimizing crop predictions and planning.

Chapter 3

System Design

The Integrated Agricultural Decision Support System (IADSS) is a comprehensive solution engineered to address the challenges faced by modern farmers. It revolutionizes modern farming practices by seamlessly integrating historical data analysis, soil attribute evaluation, machine learning algorithms, advanced meteorological data, GPS-based location services, and a user-friendly interface. This holistic solution empowers farmers with precise crop predictions, personalized fertilizer recommendations, real-time weather forecasts, and intuitive features for seamless accessibility.Below are the detailed information about the integrated modules in the project:

1. Crop Prediction and Recommendation IADSS employs historical data analysis and soil attribute evaluation to predict suitable crops for cultivation. By considering soil composition, climate conditions, and past crop performance, the system recommends optimal crops for a given location. This feature aids farmers in making informed decisions regarding crop choices aligned with their environmental conditions.

2. Fertilizer Recommendation Based on the predicted crop, IADSS provides tailored fertilizer recommendations. Leveraging machine learning algorithms and soil test inputs, the system suggests specific types and quantities of fertilizers optimized for the selected crop and soil conditions. This personalized advice aims to enhance crop productivity while minimizing environmental impact through judicious fertilizer use.

3. Real-time Weather Forecasting The system integrates advanced meteorological data sources to deliver precise, location-specific weather forecasts. By leveraging cutting-edge algorithms, IADSS provides farmers with real-time information on temperature, rainfall, humidity, and wind patterns. This data empowers

farmers to schedule agricultural activities, such as planting and irrigation, according to forecasted weather conditions, minimizing risks associated with adverse weather.

4. GPS-Based Location Services IADSS utilizes GPS functionality to determine the user's location automatically. When enabled, the system fetches local weather information, dynamically updating the user interface to display current weather conditions in the background. This feature ensures users receive tailored weather updates based on their specific geographical location.

5. User-Friendly Interface The system offers an intuitive and interactive dashboard accessible via smartphones, tablets, and computers. The user interface is designed to accommodate farmers with varying levels of technological proficiency. It includes search functionalities for manual location input and seamless navigation across features.

3.1 UML Diagrams

Unified Modeling Language (UML) is a standardized visual modeling language widely used in software engineering and system development. UML diagrams are graphical representations that help in understanding, designing, and documenting various aspects of a system. UML provides a rich set of diagram types, each serving a specific purpose. Some of the commonly used UML diagrams include use case diagrams, class diagrams, sequence diagrams, activity diagrams, and many more. UML diagrams serve as a common language for communication between stakeholders, analysts, designers, and developers, facilitating the understanding and visualization of complex systems. They aid in requirements analysis, system design, and implementation. Overall, UML diagrams provide a powerful and standardized means to represent different aspects of a system and promote effective software development practices

3.1.1 Use Case Diagram

A Use Case Diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-Case analysis. Use case diagrams in UML are visual representations that capture the functional requirements of a system from the perspective of its users or external entities. They provide an overview of the system's functionalities and showcase the interactions between users, known as actors, and the system itself. Actors represent the roles that users or external entities play when interacting with the system and are depicted as stick figures or icons outside the system boundary. Use cases, depicted as ovals or ellipses within the system boundary, represent specific actions or functionalities that the system performs to provide value to its actors. Relationships such as associations, generalizations, and dependencies illustrate the connections and dependencies between actors and use cases. Use case diagrams help stakeholders understand the system's functionalities, identify the actors involved, and visualize how users interact with the system to achieve their goals. They serve as a foundation for requirements gathering, system design, and effective communication between project teams.Use-case diagram for our project is shown in fig 3.1.

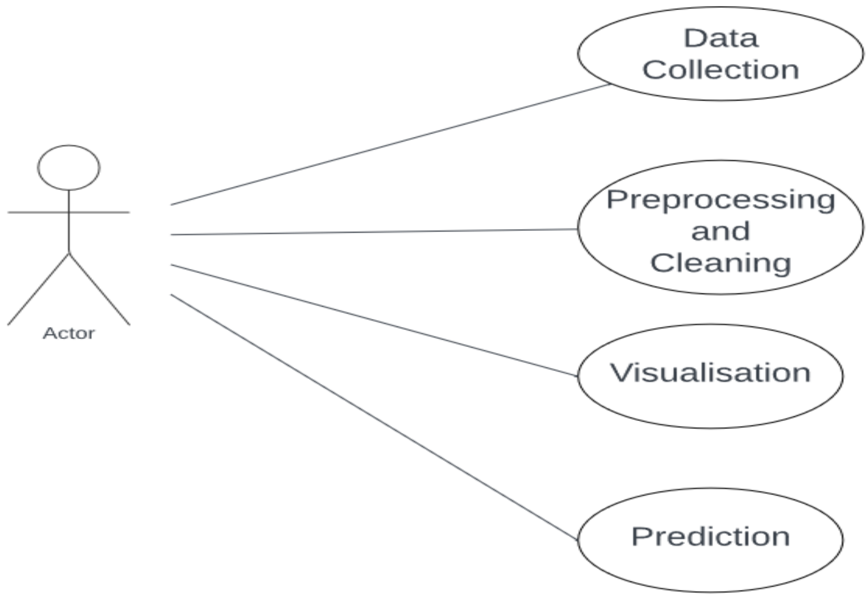

Figure 3.1: Use Case Diagram

3.1.2 Sequence Diagram

The sequence diagram represents the flow of messages in the system and is also termed as an event diagram. It helps in envisioning several dynamic scenarios. It portrays the communication between any two lifelines as a time-ordered sequence of events, such that these lifelines took part at the run time. In UML, the lifeline is represented by a vertical bar, whereas the message flow is represented by a vertical dotted line that extends across the bottom of the page. It incorporates the iterations as well as branching.Sequence Diagrams captures the interaction that takes place in a collaboration that either realizes a use case or an operation (instance diagrams or generic diagrams), high-level interactions between user of the system and the system, between the system and other systems, or between subsystems (sometimes known as system sequence diagrams).Sequence diagram for our project is shown in fig 3.2.

Figure 3.2: Sequence Diagram

3.2 Tools and Libraries used

The following are the tools and libraries involved in the project :

- **Google Colab :** While not a specific Python module, Google Colab is a cloud-based environment utilized here to run Python code interactively, providing access to GPUs and facilitating collaborative coding.
- **Pandas :** Pandas is employed for data manipulation and handling, especially for reading and managing structured data from CSV files.
- **NumPy :** NumPy is essential for numerical computations, especially when dealing with arrays and mathematical operations on large datasets.
- **Matplotlib and Seaborn :** These libraries are used for data visualization, helping in the visualization of images, graphs, and statistical plots.
- **Sklearn:** Scikit-Learn, also known as sklearn is a python library to implement machine learning models and statistical modelling. Through scikit-learn, we can implement various machine learning models for regression, classification, clustering, and statistical tools for analyzing these models.
- **Flask:** Python's Flask micro web framework is well-liked and frequently used to create online apps. It offers a straightforward and adaptable method for developing Python-based web applications and APIs (Application Programming Interfaces).

Chapter 4

Methodology

4.1 Proposed System

The proposed Integrated Agricultural Decision Support System (IADSS) helps farmers with crop selection, fertilizer recommendation, and weather forecasting. IADSS uses historical and current data on soil, weather, and crops to train a machine learning model called random forest, which predicts the best crop and fertilizer for a given location and crop,also uses an external API to get real-time weather information for the user's location, which is detected by GPS or manual input. IADSS aims to improve agricultural productivity and sustainability by providing farmers with precise and personalized insights.

The architecture 4.1 of IADSS reflects a holistic approach, combining agriculture, machine learning algorithms and web concepts to create an effective and user-friendly solution for enhancing resilience and prosperity of farming communities on a global scale.

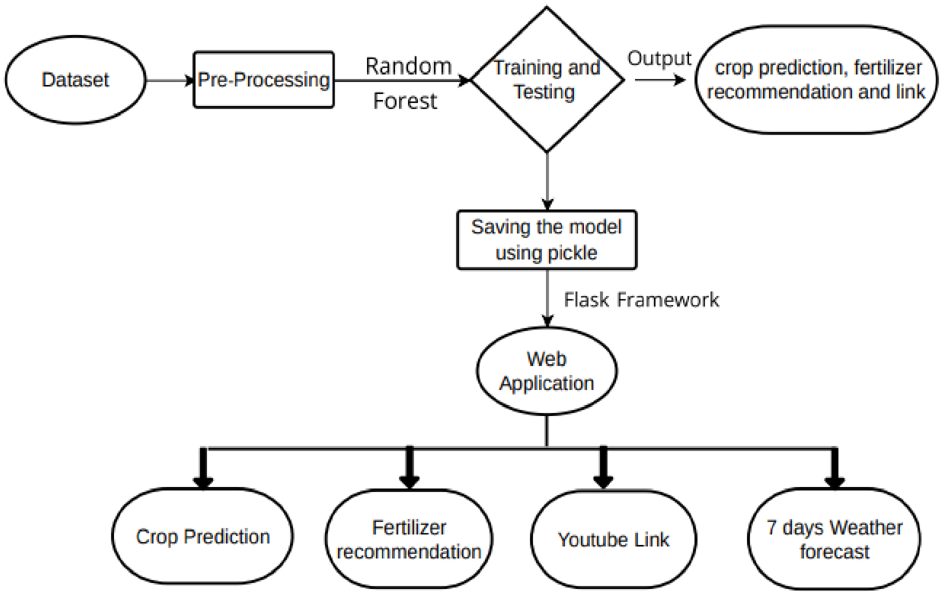

Figure 4.1: Architecture

4.2 Dataset

The Dataset is a publicly available dataset containing 4514 variations that aims at providing a clear picture of Crops and Fertilizer to be used to increase productivity. The features of dataset are: District_Name, Soil_color, Nitrogen, Phosphorus, Potassium, pH, Rainfall, Temperature, Crop, Fertilizer, Link.This data frame consists of eleven factors in which District_Name, Soil_color, Nitrogen, Phosphorus, Potassium, pH, Rainfall, Temperature are input parameters and Crop, Fertilizer, Link are output target values, which are trained by various machine learning techniques to predict the crop and the fertilizer which should be used for better yield of a crop and a link which provides details about the crop. Dataset is taken from Kaggle named Crop and Fertilizer and few modifications were done based on our requirements and is shown in fig 4.2 .The detailed description of dataset is given in table 4.1.

P7

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	District_Na	Soil_color	Nitrogen	Phosphoru	Potassium	pH	Rainfall	Temperatu	Crop	Fertilizer	Link		
2	Hyderaba	Black	75	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
3	Hyderaba	Black	80	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
4	Hyderaba	Black	85	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
5	Hyderaba	Black	90	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
6	Hyderaba	Black	95	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
7	Hyderaba	Black	100	50	100	6.5	1000	20	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
8	Hyderaba	Black	75	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
9	Hyderaba	Black	80	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
10	Hyderaba	Black	85	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
11	Hyderaba	Black	90	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
12	Hyderaba	Black	95	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
13	Hyderaba	Black	100	55	105	7	1100	25	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
14	Hyderaba	Black	75	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
15	Hyderaba	Black	80	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
16	Hyderaba	Black	85	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
17	Hyderaba	Black	90	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
18	Hyderaba	Black	95	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
19	Hyderaba	Black	100	60	110	7.5	1200	30	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
20	Hyderaba	Black	75	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
21	Hyderaba	Black	80	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
22	Hyderaba	Black	85	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
23	Hyderaba	Black	90	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
24	Hyderaba	Black	95	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
25	Hyderaba	Black	100	50	115	6.5	1300	35	Sugarcane	Urea	https://youtu.be/2t5Am0xLTOo		
26	Hyderaba	Black	75	55	100	7	1400	20	Sugarcane	DAP	https://youtu.be/2t5Am0xLTOo		
27	Hyderaba	Black	80	55	100	7	1400	20	Sugarcane	DAP	https://youtu.be/2t5Am0xLTOo		
28	Hyderaba	Black	85	55	100	7	1400	20	Sugarcane	DAP	https://youtu.be/2t5Am0xLTOo		

Figure 4.2: Attributes of the Dataset

S.No	Feature	Description
1	District_Name	The name of the district where the data was collected.
2	Soil_color	Color or type of soil.
3	Nitrogen	Quantity of nitrogen present in the soil (measurement unit is not specified).
4	Phosphorus	Quantity of phosphorus in the soil.
5	Potassium	Quantity of potassium in the soil.
6	pH	pH level of the soil.
7	Rainfall	Amount of rainfall in the region (measurement unit is not specified generally in mm).
8	Temperature	Temperature in the region (measurement unit is not specified generally in Celcius).
9	Crop	Type of crop being cultivated.
10	Fertilizer	Type of fertilizer recommended for the crop.
11	Link	Link to a video or resource related to the crop (YouTube link provided).

Table 4.1: Features and their description

4.3 Data Pre-processing

In this stage the data set is pre-processed i.e handling missing data, null values, duplaicate values, then Encoding of features named District_Name & Soil_color is done through Label Encoder to convert categorical values into numeric values for prediction fig 4.3 depicts the process. Later after the prediction of the result is done, values are to be converted back to their original values for next predictions. So, this back conversion to original values of features District_Name & Soil_color is done through inverse function shown in fig 4.4. This process made us train the data easily for predictions.

```
# Convert original values to encoded values
label_encoder_district = preprocessing.LabelEncoder().fit(dataset['District_Name'])
label_encoder_soil_color = preprocessing.LabelEncoder().fit(dataset['Soil_color'])

encoded_district = label_encoder_district.transform([district])[0]
encoded_soil_color = label_encoder_soil_color.transform([soil_color])[0]

input_data = pd.DataFrame(
    [[encoded_district, encoded_soil_color, nitrogen, phosphorus, potassium, pH, rainfall, temperature]],
    columns=['District_Name', 'Soil_color', 'Nitrogen', 'Phosphorus', 'Potassium', 'pH', 'Rainfall', 'Temperature']
)
# Convert the encoded values back to original values for prediction
predicted_district_encoded = label_encoder_district.transform([district])[0]
predicted_soil_color_encoded = label_encoder_soil_color.transform([soil_color])[0]
# Transform categorical data for the entire dataset
dataset_encoded = dataset.copy()
dataset_encoded['District_Name'] = label_encoder_district.transform(dataset_encoded['District_Name'])
dataset_encoded['Soil_color'] = label_encoder_soil_color.transform(dataset_encoded['Soil_color'])
# Split the dataset into training and testing sets
```

Figure 4.3: Encoding of attributes

```
# Convert the predicted encoded values back to original values
original_predicted_district = label_encoder_district.inverse_transform([predicted_district_encoded])[0]
original_predicted_soil_color = label_encoder_soil_color.inverse_transform([predicted_soil_color_encoded])[0]
```

Figure 4.4: Inverse Encoding of attributes

4.4 Training and Testing

After Pre-Processing phase, Dataset is splitted into two halves - Training and Testing sets for predictions.So, For splitting the dataset into training set and test set we used package sklearn model selection and import train_test_split, this helps to successfully split the training data and test data according to the given split value. The split size of the test data should be 20% and training data should be 80%. Dataset is trained using various machine learning algorithms such as Decision Trees, Random Forests, SVM, and Naive Bayes.

4.4.1 Random Forest Classifier

Random Forest is an ensemble learning method known for its high accuracy in both classification and regression tasks. By constructing multiple decision trees during training and aggregating their predictions, it effectively minimizes overfitting. This algorithm is robust and excels in handling large datasets with diverse features, requiring minimal preprocessing efforts. It not only evaluates feature importance for better interpretability but also efficiently manages unbalanced data, providing balanced class probabilities. Its versatility is evident in applications ranging from customer churn prediction to bioinformatics. The advantages of high accuracy, resistance to overfitting, large dataset handling, feature importance assessment, and efficiency with unbalanced data make Random Forest a powerful and widely-used tool in the machine learning landscape.Structure of Random Forest classifier model is shown in fig 4.5 .

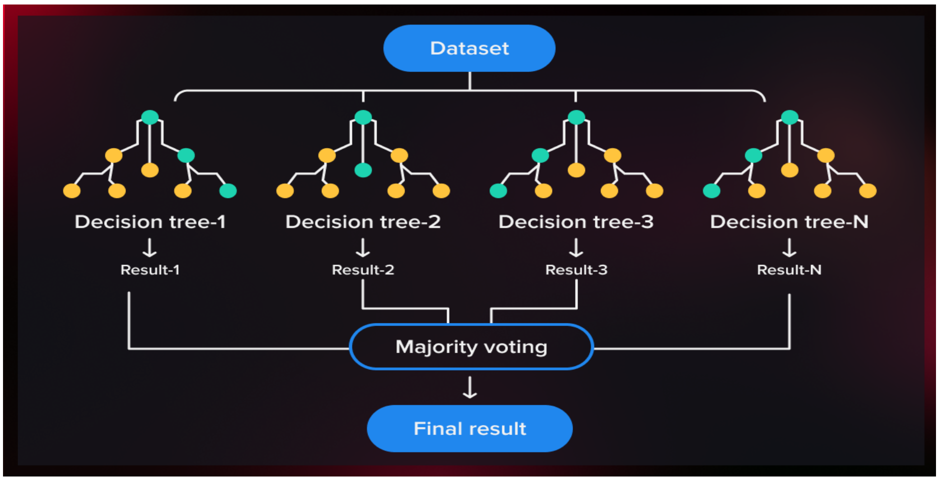

Figure 4.5: Structure of Random Forest

4.4.2 Decision Tree Classifier

Decision Tree is a powerful supervised learning tool adept at handling both classification and regression challenges, though it's often favored for classification. Its intuitive tree structure, featuring decision nodes for dataset features and leaves for outcomes, ensures easy interpretability, making it accessible to non-experts. Model Defnition of this classifier is represented in fig 4.6. Noteworthy advantages include its ability to seamlessly handle both numerical and categorical data, requiring minimal preprocessing, and its inherent feature selection capability, aiding in identifying crucial factors influencing decisions. Decision Trees find versatile applications in medical diagnosis, marketing analytics, fault diagnosis systems, and environmental sciences, demonstrating their adaptability and effectiveness across domains.

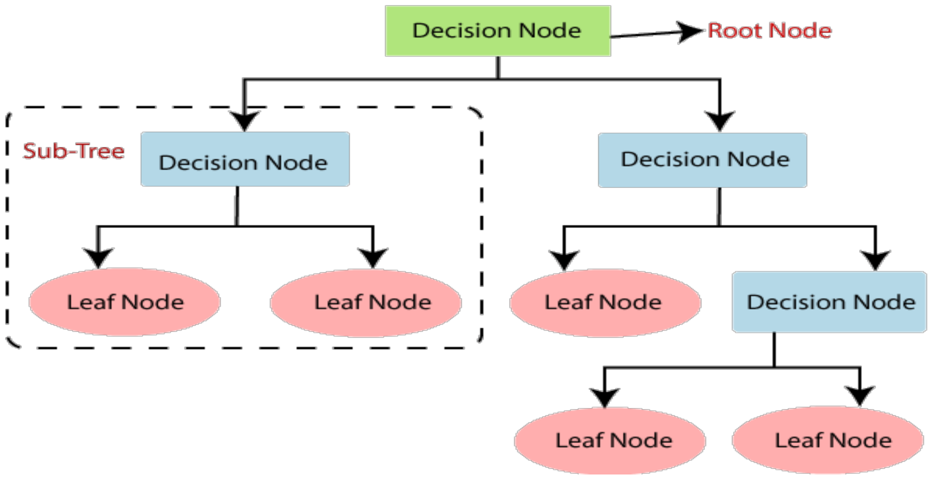

Figure 4.6: Model Defination of DecisionTree Classifier

4.4.3 Support vector machine

Support Vector Machine (SVM) is a widely used supervised learning algorithm, primarily applied to classification tasks in machine learning. Its objective is to create an optimal decision boundary, known as a hyperplane, that effectively segregates n-dimensional space into distinct classes. SVM identifies crucial extreme points, termed support vectors, to construct this hyperplane, algorithm structure is represented in fig 4.7. The algorithm has two main types: Linear SVM, suitable for linearly separable data classifiable by a straight line, and Non-linear SVM, designed for data that requires a more complex, non-linear decision boundary. In essence, SVM excels in creating robust decision boundaries for effective classification in both linearly and non-linearly separable datasets.

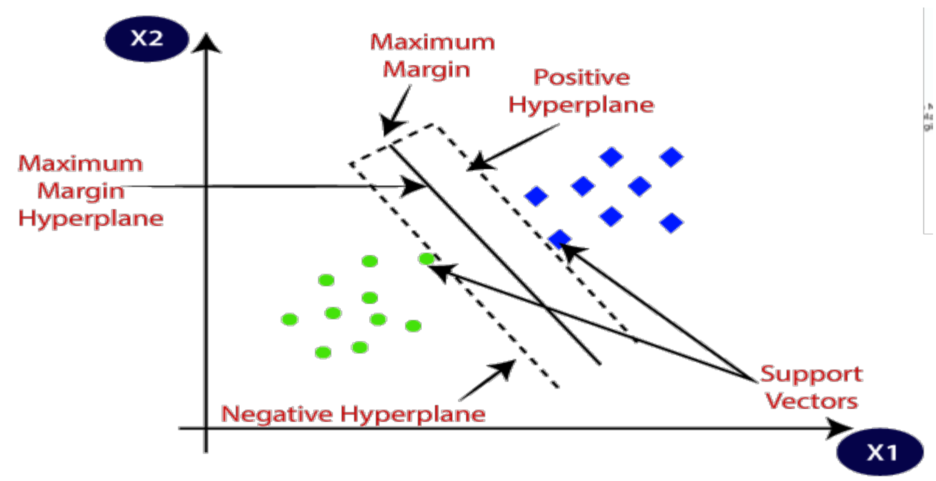

Figure 4.7: Structure of Support vector algorithm

4.4.4 Naive Bayes Classifer

Naïve Bayes is a supervised learning algorithm rooted in Bayes' theorem, predominantly employed for solving classification problems, especially in text classification with high-dimensional datasets as shown in fig 4.8. Recognized for its simplicity and effectiveness, Naïve Bayes is a probabilistic classifier, making predictions based on the probability of an object belonging to a certain class. It finds application in diverse tasks such as spam filtration, sentiment analysis, and article classification. Advantages include its speed, simplicity, applicability to both binary and multi-class classifications, and superior performance in multi-class predictions. However, its main drawback lies in the assumption of feature independence, limiting its ability to learn relationships between features. In essence, Naïve Bayes is a fast and efficient choice for classification tasks, particularly well-suited for text-related applications.

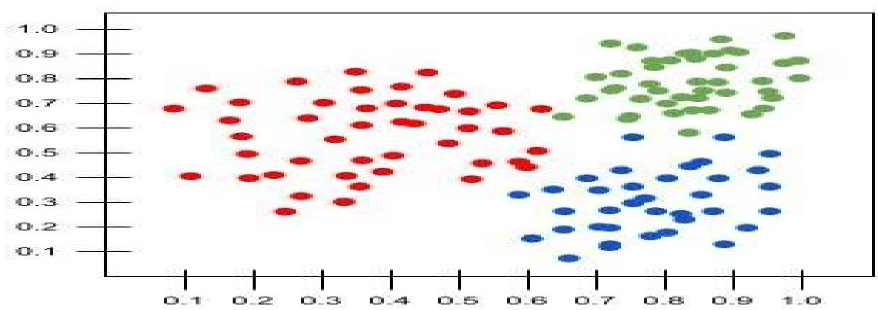

Figure 4.8: Structure of Naive Bayes Algorithm

We trained the machine learning model using various algorithms such as Decision Trees, Random Forests, SVM, and Naive Bayes. After rigorous testing, we found Random Forest and Decision Tree to be more effective among them. Then we used Random Forest and 'pickle' to save and integrate the model into the web application.The following figures 4.9, 4.10 represents the importing, training of ml model:

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import GaussianNB

from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
```

Figure 4.9: importing All Algorithms classifiers

```
X_train, X_test, y_train, y_test = train_test_split(
    dataset_encoded.drop(['Crop', 'Fertilizer', 'Link'], axis=1),
    dataset_encoded['Crop'],
    test_size=0.2,
    random_state=42
)
# Train the random forest model
model_crop = RandomForestClassifier(n_estimators=100, random_state=42)
model_crop.fit(X_train, y_train)

model_filename = '/content/crop_prediction_model.pkl'
with open(model_filename, 'wb') as model_file:
    pickle.dump(model_crop, model_file)

# Save the label encoders
label_encoder_district_filename = '/content/label_encoder_district.pkl'
label_encoder_soil_color_filename = '/content/label_encoder_soil_color.pkl'

with open(label_encoder_district_filename, 'wb') as encoder_file:
    pickle.dump(label_encoder_district, encoder_file)

with open(label_encoder_soil_color_filename, 'wb') as encoder_file:
    pickle.dump(label_encoder_soil_color, encoder_file)

# Make predictions
predicted_crop_encoded = model_crop.predict(input_data)
```

Figure 4.10: Training and Saving the model

4.5 Prediction

After taking input values i.e soil parameters , loaction specific weather conditions provided by user the model predicts the suitable crop, fertilizer and also a link about crop details shown in fig 4.11.

Figure 4.11: Output for given input values in trained model

Chapter 5

Implemention

After Predictions were made, then we started developing the user-freindly web application using Flak Framework by deploying the saved trained model which was done through pickle along with the label_encoders files. It is integrated with the model and also OpenWeatherMap for 7 -days Weather forecast module which is automatic gps and location specific based with dynamic background change . When User inputs the values of soil parameters and location-specific weather conditions the model predicts the crop based on the inputs with image , and recommends crop-specific fertilizer with image to user along with a link of the crop details integrated with a 7-day weather forecast with dynamic background changes. The following figures represents the inputs and outputs of the application . Fig 5.1 shows the user input format.For the given input the crop predicted i.e wheat is shown in fig 5.2 and fertilizer recommended i.e npk is shown in fig 5.3.images 5.4 and 5.5 are the output predictions of another set of input parameters.

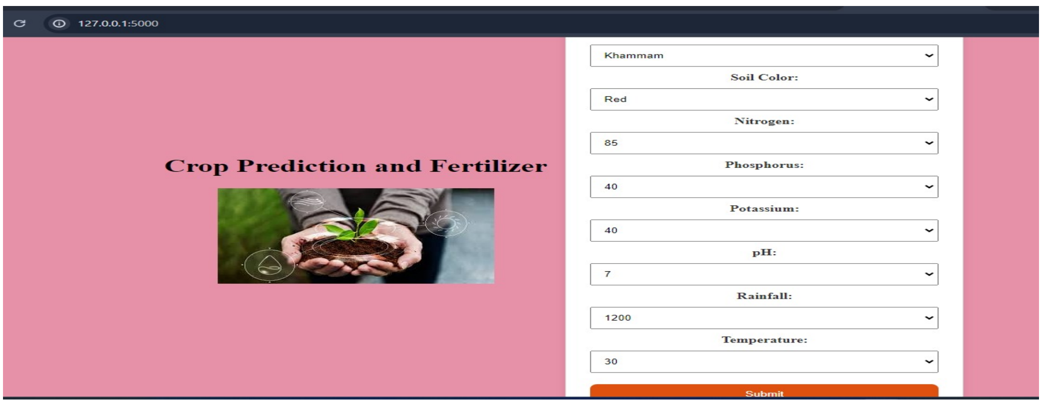

Figure 5.1: Web Output taking input values

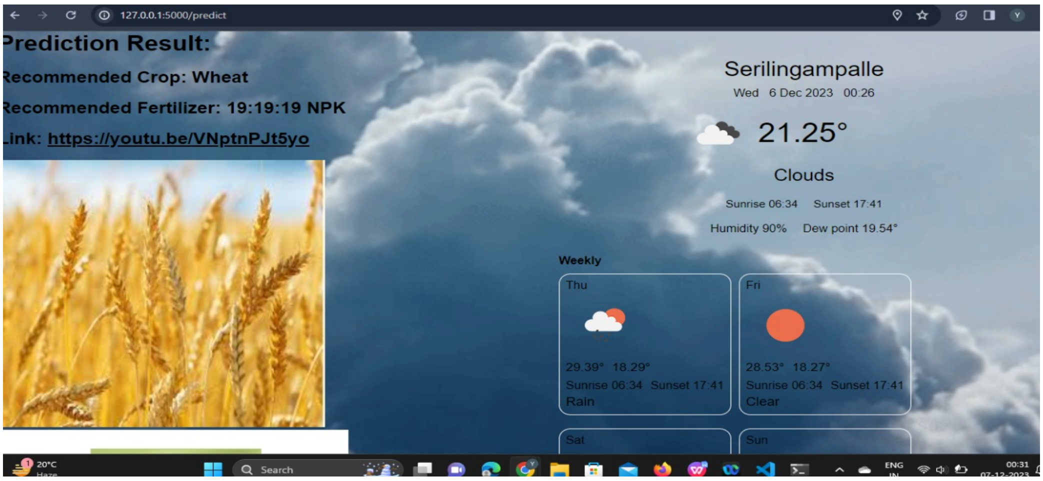

Figure 5.2: Output Representing Wheat Crop

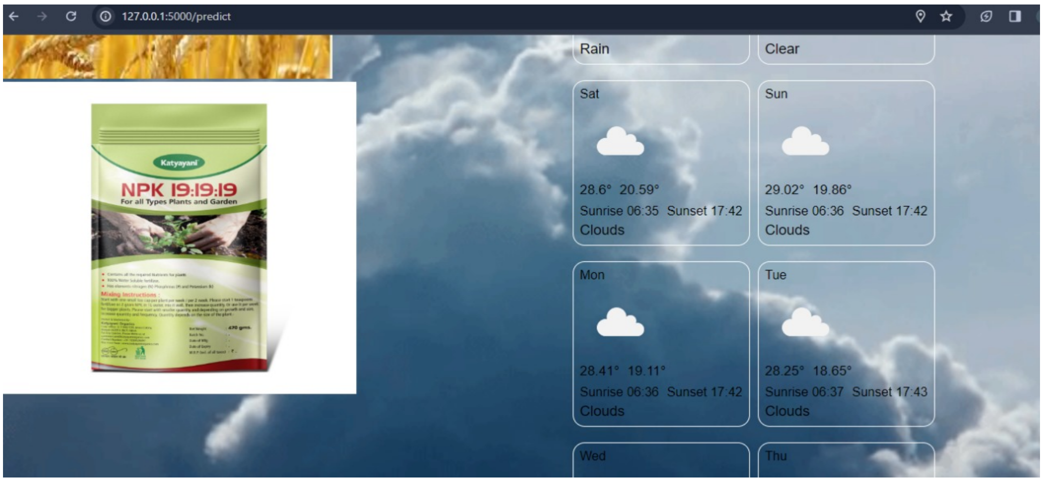

Figure 5.3: Output Representing NPK Fertilizer

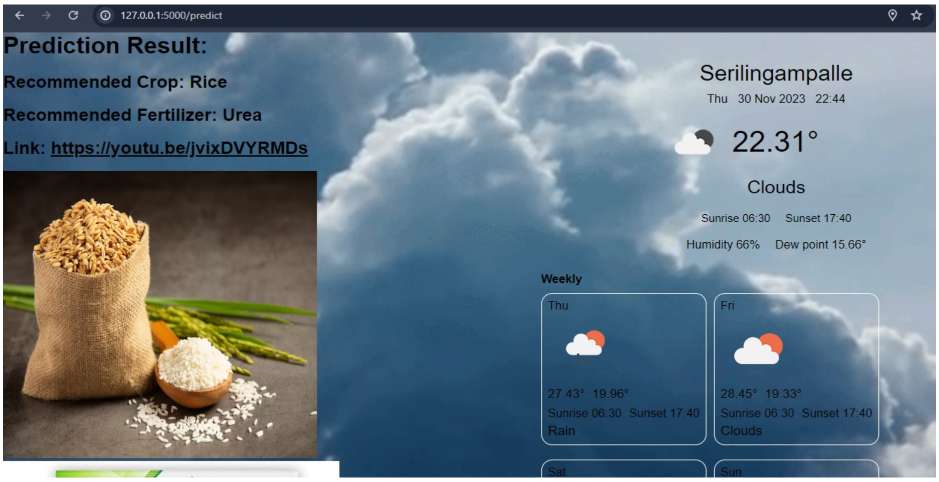

Figure 5.4: Recommended crop and geographical location weather forecast

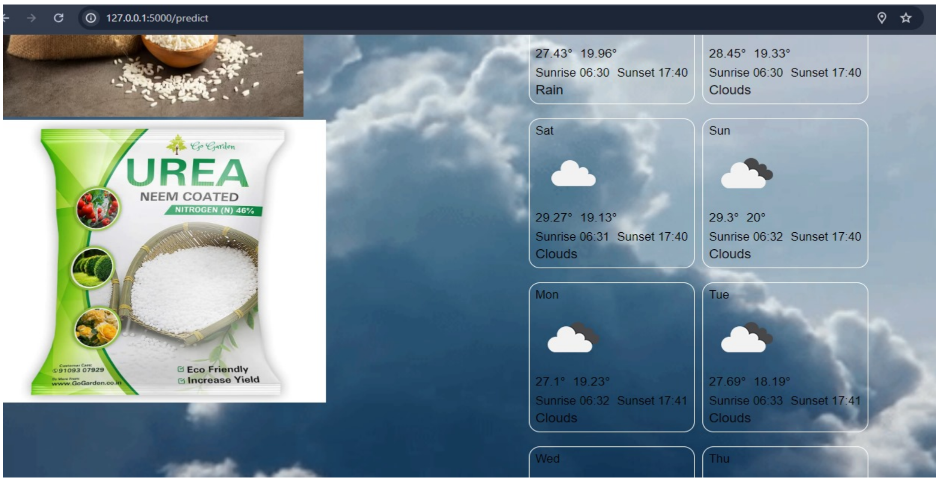

Figure 5.5: Recommended fertilizer for crop predicted

Chapter 6

Results and Discussions

We trained the model using Four prominent machine learning algorithms – Random Forest, Decision Tree, Support Vector Machine (SVM), and Naive Bayes – are employed. These algorithms are evaluated based on key performance metrics, including accuracy, F1 score, recall, and precision. Among all Random Forest Classifier and Decision Tree Classifier have higher accuracy and performance compare to other algorithms. After rigorous testing, we found Random Forest to be the most effective, and we used 'pickle' to save and integrate the model into a web application, which predicts both crop and fertilizer based on soil parameters and weather conditions and provides a link of crop information ,with integration of weather forecast. This ensures farmers have a reliable tool for informed crop decisions based on real-time conditions.Fig 6.1 represents the graphical representation of all algorithms comparision. Fig 6.2 represents the comparision of four algorithms. Table 6.1 represents the metrics values of all algorithms

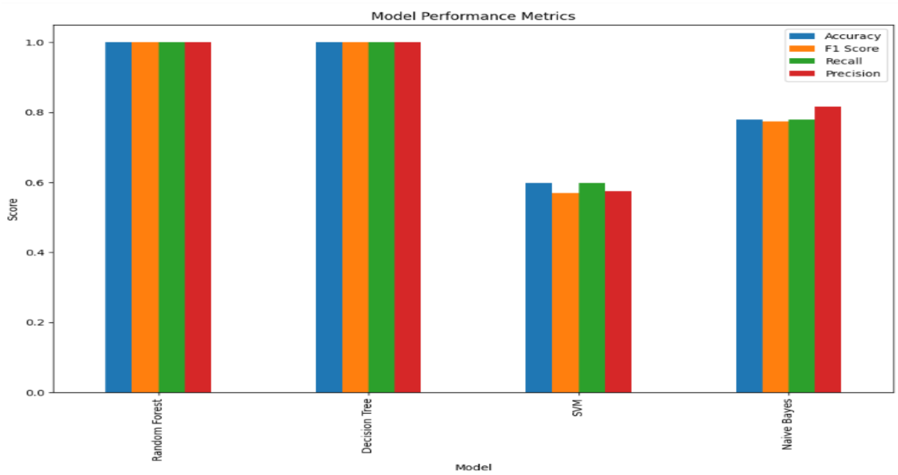

Figure 6.1: Comparision chart of all algorithms based on metrics

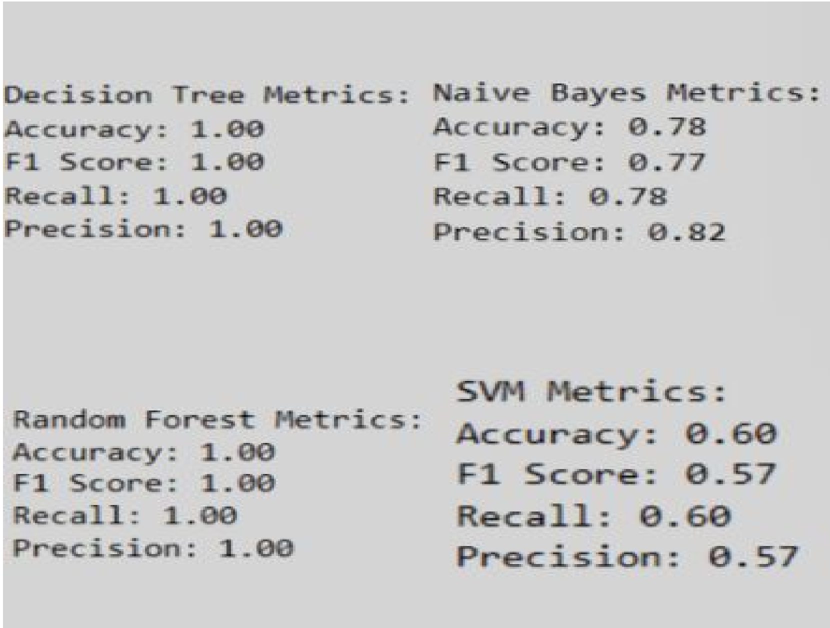

Figure 6.2: Comparison of all algorithms based on metrics

Algorithm	**Accuracy**	**F1 Score**	**Precision**	**Recall**
Random Forest	1	1	1	1
Decision Tree	1	1	1	1
SVM	0.599114	0.568678	0.599114	0.57415
Naive Bayes	0.779623	0.773655	0.779623	0.81643

Table 6.1: Metrics Values for Different Algorithms

Chapter 7

Conclusion and Future Scope

The Integrated Agricultural Decision Support System (IADSS) represents a transformative approach aimed at revolutionizing traditional agricultural methodologies. Its primary focus on data-driven decision-making underscores a pivotal shift in the industry, harnessing technological advancements to optimize agricultural productivity, sustainability, and resilience. Through meticulous analysis of agricultural data, IADSS endeavors to elevate productivity by fine-tuning resource allocation, crop management strategies, and identifying optimal timelines for planting and harvesting. Moreover, its emphasis on sustainability underscores a commitment to eco-friendly practices, seeking to minimize environmental impact while enhancing crop resilience amidst the challenges posed by shifting climatic conditions.

”In the future, IADSS envisions significant advancements, focusing on enhancing predictive capabilities, leveraging advanced data analytics, and integrating cutting-edge IoT devices for real-time monitoring. These developments aim to provide more accurate forecasts, covering various crucial agricultural factors such as weather patterns, pest outbreaks, crop yields, and market demands. The integration of advanced data analytics and IoT devices seeks to optimize decision-making for farmers, fostering efficiency, sustainability, and resilience in the agricultural sector globally.”

Bibliography

[1] S. Paudel, R. Nakarmi, P. Giri, S. B. Karki, *et al.*, "Prediction of crop yield based-on soil moisture using machine learning algorithms," in *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, pp. 491–495, IEEE, 2022.

[2] S. M. Pande, P. K. Ramesh, A. ANMOL, B. Aishwarya, K. ROHILLA, and K. SHAURYA, "Crop recommender system using machine learning approach," in *2021 5th international conference on computing methodologies and communication (ICCMC)*, pp. 1066–1071, IEEE, 2021.

[3] T. Setiadi, F. Noviyanto, H. Hardianto, A. Tarmuji, A. Fadlil, and M. Wibowo, "Implementation of naïve bayes method in food crops planting recommendation," *Int. J. Sci. Technol. Res*, vol. 9, no. 02, pp. 4750–4755, 2020.

[4] S. R. Sani, S. V. S. Ummadi, S. Thota, N. Muthineni, V. S. S. Swargam, and T. S. Ravella, "Crop recommendation system using random forest algorithm in machine learning," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pp. 501–505, IEEE, 2023.

[5] M. Shahhosseini, G. Hu, I. Huber, and S. V. Archontoulis, "Coupling machine learning and crop modeling improves crop yield prediction in the us corn belt," *Scientific reports*, vol. 11, no. 1, p. 1606, 2021.

[6] P. Rajak, A. Ganguly, S. Adhikary, and S. Bhattacharya, "Internet of things and smart sensors in agriculture: Scopes and challenges," *Journal of Agriculture and Food Research*, vol. 14, p. 100776, 2023.

[7] T. Mahesh *et al.*, "Prediction of crop yield based-on soil moisture using machine learning algorithms," *International Journal of Information Technology, Research and Applications*, vol. 2, no. 1, pp. 33–41, 2023.

[8] NikhilYadav and et al., "Crop yield and fertilizers prediction using decision tree algorithm," *International Journal of Engineering Applied Sciences and Technology (IJEAST)*, vol. 05, pp. 187–193, 2020.

[9] LakshmiDevi and et al., "The review on crop and fertilizer prediction using deep learning," *International Journal of Research Publication and Reviews (IJRPR)*, vol. 03, pp. 180–186, 2022.

[10] DhruviGosai and et al., "The review on crop and fertilizer prediction using deep learning," *International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCEIT)*, vol. 07, pp. 554–557, 2021.