REAL-TIME CROP PREDICTION AND FERTILIZER RECOMMENDATION SYSTEM USING MACHINE LEARNING AND IOT

The Project submitted to

Sant Gadgebaba Amravati University, Amravati

Towards partial fulfillment of the Degree of

Bachelor of Engineering

In

Information Technology

Guided by
Prof. S. D. Padiya

Submitted by
Mr. Atharva Raut
Mr. Bhavy Mittal
Mr. Mayur Patel
Mr. Vedant Polshettiwar

DEPARTMENT OF INFORMATION TECHNOLOGY
SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING, SHEGAON (M.S.)

2022- 2023

12

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING, SHEGAON

EST.1983

2022-2023

CERTIFICATE

This is to certify that **Mr. Atharva Raut, Mr. Bhavy Mittal, Mr. Mayur Patel, Mr. Vedant Polshettiwar** students of final year B.E. (Information Technology) in the year 2022-2023 of the Information Technology Department of this institute have completed the project work entitled **"Real-time crop prediction and fertilizer recommendation system using machine learning and IoT"** based on syllabus and has submitted a satisfactory account of his/her work in this report which is recommended for the partial fulfilment of the degree of Bachelor of Engineering in Information Technology.

Prof. S. D. Padiya
(Project Guide)

Dr. A S Manekar
Head of the Department
SSGMCE, Shegaon

Dr. S. B. Somani
Principal
SSGMCE, Shegaon

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING, SHEGAON

2022-2023

<u>CERTIFICATE</u>

This is to certify that the project work entitled **"Real-time crop prediction and fertilizer recommendation system using machine learning and IoT"** submitted by **Mr. Atharva Raut, Mr. Bhavy Mittal, Mr. Mayur Patel, Mr. Vedant Polshettiwar** students of final year B.E. (Information Technology) in the year 2022-2023 of the Information Technology Department of this institute, is a satisfactory account of his work based on the syllabus which is approved for the award of the degree of Bachelor of Engineering in Information Technology.

Internal Examiner	External Examiner
Date:	Date:

ACKNOWLEDGEMENT

The real spirit of achieving goals through the way of excellence and lustrous discipline. We would have never succeeded in completing our task without the cooperation, encouragement and help provided to us by various personalities.

We would like to take this opportunity to express our heartfelt thanks to our guide ***Prof. S. D. Padiya*** *for his esteemed guidance and encouragement, especially through difficult times. His suggestions broaden our vision and guide us to succeed in this work. We are also very grateful for his guidance and comments while studying part of our project and learning many things under his leadership.*

We would also like to sincerely thank ***Prof.F. I. Khandwani****, Project-In-Charge for his valuable support and feedback during the entire course of the project.*

We also extend our thanks to ***Dr. A. S. Manekar****, Head of Information Technology Department, Shri Sant Gajanan Maharaj College of Engineering, Shegaon for providing us with a variety of opportunities and inspirations to gather professional knowledge and material that made us consistent performers.*

We also extend our thanks to ***Dr. S. B. Somani****, Principal, Shri Sant Gajanan Maharaj College of Engineering, Shegaon for providing us the infrastructure and facilities without which it was impossible to complete this work.*

Also, we would like to thank all teaching and non-teaching staff of the department for their encouragement, cooperation and help. Our greatest thanks to all those who wished us success, especially parents and friends.

Student Names

1. ***Atharva Raut***
2. ***Bhavy Mittal***
3. ***Mayur Patel***
4. ***Vedant Polshettiwar***

ABSTRACT

The common issue of farmers is they are not aware of the crop which suits their soil quality, and soil nutrients, and to check the soil quality only way is to get a sample to the lab and get tested to predict the crop is suitable for cultivation according to their soil type, also the fertilizers are in various varieties that it is hard to choose the best for the crop but it is a costly and time-consuming process hence our project will try to solve this issue using IoT and ML. The system will help farmers cultivate proper crops for better yield production. This project will focus mainly on predicting the yield of the crop by applying various machine learning techniques on real-time data from sensors or input which will be given by the user or farmer also later suggest suitable fertilizers on the need of soil. The classifier models used here include Naïve Bayes and Random Forest. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to induce the most yield by considering factors like temperature, rainfall, area, etc.

As a result, the system will predict the above 80% best crop for given data of soil and nutrients, temperature and recommend fertilizers suitable for crop and soil using machine learning techniques from which farmers can grow their production.

Keywords *– crop prediction, soil moisture sensor, soil npk sensor, random forest, crop yield*

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE NO.
1.	**INTRODUCTION**	1
	1.1 Preface	1
	1.2 Statement of problem	3
	1.3 Objectives of Project	4
	1.4 Scope and Limitations of the Project	5
	1.5 Organization of the Project	6
2.	**LITERATURE SURVEY**	7
3.	**ANALYSIS**	29
	3.1 Detailed Statement of the Problem	29
	3.2 Requirement Specifications	30
	3.2.1 Software Requirement	30
	3.2.2 Hardware Requirement	31
	3.3 Functional Requirement	31
	3.3.1 Data Flow Diagram	31
	3.4 Non Functional Requirement	33
	3.5 Feasibility Study	34
	3.6 Use Case Diagrams	35
	3.7 Use Case Specification	37
4.	**DESIGN**	38
	4.1 Design goals	38
	4.2 Design Strategy	39
	4.3 Architecture Diagram	41
	4.4 Class Diagram	42
	4.5 Sequence Diagram	43
	4.6 Collaboration Diagram	45
	4.7 State Chart Diagram	46
	4.8 Activity Diagram	48
5.	**IMPLEMENTATION**	50
	5.1 Implementation Strategy	50

5.2 Hardware Platform Used 51
5.3 Software Platform Used 51
5.4 Hardware Specification 51
5.4.1 Arduino Uno 52
5.4.2 Soil NPK Sensor 54
5.4.3 Soli Moisture Sensor 59
5.5 Software Specification 63
5.5.1 Arduino IDE 64
5.5.2 Python libraries Used 64
5.6 Deployment Diagram 70
5.7 Implementation Level Details 71
5.7.1 Dataset Details 71
5.7.2 Module 2: Preprocessing 73
5.7.3 Module 3: Random Forest Model 75
5.8 Testing 79
6. CONCLUSION 80
FUTURE WORK 81
User Manual 82
REFERENCES 93
Dissemination of Work 98
Source code Listings 101
Information of Members 102

LIST OF FIGURES

Sr. No.	Figure Name	Page No.
3.1	DFD Level 1	31
3.2	DFD level 2	32
3.3	Use case diagram	36
4.1	Strategy Diagram	39
4.2	Architecture Diagram	41
4.3	Class Diagram	43
4.4	Sequence Diagram	44
4.5	Collaboration Diagram	46
4.6	State Chart Diagram	47
4.7	Activity Diagram	49
5.1	Pin diagram of Arduio Uno	53
5.2	Importance of Nutrients	54
5.3	Soil NPK sensor	55
5.4	Probes of NPK sensor	56
5.5	NPK Sensor Pin diagram	57
5.6	RS-485 transceiver module	58
5.7	Circuit diagram for NPK sensor	59
5.8	Soil Moisture Sensor	60
5.9	Probes	62
5.10	Moisture sensor module	62
5.11	Soil Moisture Sensor Circuit	63
5.12	Deployment Diagram	70
5.13	Implementation Model	71
5.14	Dataset	72
5.15	Pre-Processing	74
5.16	Prediction Model	76
5.17	Random Forest model simplified	77
5.18	Accuracy Comparison of Models	78

1. INTRODUCTION

1.1 Preface

In India, there is a wide range of both physical and cultural diversity. Nearly everyone in India is reliant on agricultural and agriculture-related jobs. Agriculture is benefiting greatly from the Internet of Things (IoT). Farmers from a variety of issues and allow them to concentrate on other relevant careers. we believe that precision farming is one of the greatest innovations ever made. When it comes to illness prevention and early disease detection, it's doing all it can loss and crop recommendations, as well as information on weather conditions, are supplied in addition to this. Exporting his goods and assisting in the preservation of the field His actions are controlled by sensors and actuators tasks like irrigation and pesticide application should be scaled appropriately. By using all of these methods, he may maximize his profits. Keep a close eye on his area of expertise. Over half of the people in this country make their living via agriculture.[2]

According to the Economic Survey for 2016-17, Farmers in 17 states earn an average monthly wage of $1,050. Farmer suicides and diversion of Rs. 1700/-of farmland for non-farming use. In addition, 48% of farmers said they don't want their next generation to instead of taking care of their farm, the next generation should urban locations are where we wish to live. As to why the reason for this is because farmers are notoriously inept. Crop selection is an example of this picking a crop that will not provide a significant amount of income planting at the incorrect season, specific soil, and so on. It's possible that the farmer bought the property from those who have never made a choice like this before Possibly taken away from me. If you choose the wrong crop, you'll result in a lower return on investment If the family is in good health, they will be happy. If you rely on this revenue, it's quite tough to make ends meet. Correct and timely information are both readily available and easily accessible. The lack of access to current knowledge deters would-be researchers due to my experience with case studies focusing on poor countries. A mechanism has been put in place using the resources available to us. Presented a solution So this issue that will provide insights into the long-term viability of crops machine learning models are used to make suggestions taught with environmental and social considerations in mind monetary factors.[3]

Internet of Things (IoT) refers to millions of physical devices connected over the globe to collect and share information. Internet of Things is an application area which combines various technologies (software) and devices (hardware) described in. “Internet” is connectivity which helps in communication and by the term “Things” it means a combination of sensors, computing devices, smart-phones, Radio frequency identification (RFID) etc. In, the author discusses the technological advancement and cost effectiveness that has made it possible for any kind of user to use it with flexibility by making living very simple and progressively agreeable.[5]

Implementing IoT in agriculture utilizes sensors and microcontrollers to make use of the system efficiently, where sensors have the potential to get a large amount of field information. By incorporating a distinctive sensor’s data it is possible to decide on suitability and fertility of the soil. IoT helps in providing a better yield of the crop which improves productivity.

ML is an exciting application of Artificial Intelligence. It provides the ability to learn by experiences without any explicit program. The proposed model is based on simple and cost-effective hardware that can be used by agriculture officers and farmers to get good productivity of crops. SCS model is trained by classifying dataset and tested subsequently. The accuracy and performance of an ML classifier depend only on the type and size of the dataset. Our dataset used for training the model has 2200 instances for 11 crops. For dataset classification, five supervised ML algorithms (DT, SVM, KNN, RF, and NB) are used. To overcome the weaknesses of individual ML algorithm, they are ensembled for improved accuracy.[6]

Provide good user experience by having better accessibility, and usability, and providing a more efficient way for the interactions. For that, we aim to provide a web-based application that the user can use to control.

1.2 Statement of Problem

An accurate crop prediction model can help farmers to decide on what to grow and when to grow. Agricultural unpredictability, due to changing temperature season soil parameters are the factors which reduces productivity. More production should be gained with a larger population and area, however this is impossible to achieve. Farmers used to rely on word-of-mouth, but owing to climate reasons, they can no longer do so. The adequate quantity and quality of fertilizers provide the essential nutrients to the soil for the sustained production of crops. To recommend the required fertilizer that the soil need and to stop the excess use of fertilizer which are not required to save the cost, there is a need of system which can make the work of farmer easy for them.

- Solution Requirement

  We analyzed the problem statement and found the feasibility of the solution of the problem. We read a different research paper. After checking the feasibility of the problem statement. The next step is the dataset gathering and analysis. We analyzed the data set in different approaches of training like negatively or positively trained. So, after doing lot of research, we found that the balanced training of the algorithm is the best way to avoid the bias and variance in the algorithm and get a good accuracy.

- Solution Constraints

  We analyzed the solution in terms of cost, speed of processing, requirements, level of expertise, and availability of equipment's.

1.3 Objectives of the project

The objectives of the Crop Prediction and Fertilizer Recommendation System project are:

- To predict the crop yield and crop growth based on environmental factors such as temperature, humidity, and nutrient content of the soil.
- To recommend the best fertilizer for a particular crop based on the nutrient content of the soil and the crop growth stage.
- To optimize the use of fertilizers and reduce the risk of over-fertilization, which can lead to environmental pollution and damage to crops.
- To improve crop productivity and reduce crop failure by providing accurate predictions and recommendations for farmers.
- To increase the efficiency of farming practices by automating the process of crop prediction and fertilizer recommendation.
- To provide real-time data analysis and decision-making support to farmers, thereby improving the overall agriculture industry.

1.4 Scope and limitation of the Project

The project on crop prediction and fertilizer recommendation using machine learning algorithms has a wide scope in the agricultural industry. The project has the potential to contribute to sustainable agriculture, food security, and environmental conservation.

The project scope includes:

Collecting and analyzing data on soil properties, crop characteristics, weather patterns, and other relevant factors to build a comprehensive database for model training.

Developing and testing machine learning algorithms and models that can accurately predict crop yield and recommend optimal fertilizer application strategies.

Incorporating real-time data streams into the models to enhance the accuracy and timeliness of predictions and recommendations.

Limitations:

While the project has a wide scope and potential impact, there are some limitations that should be considered, including:

- Availability and quality of data: The accuracy of the machine learning models depends on the availability and quality of the data used for model training. The project may face limitations in obtaining relevant and reliable data.
- Variability in crop yield: Crop yield is influenced by various factors, including weather, pests and diseases, soil fertility, and management practices. The models may not be able to accurately predict crop yield in all situations due to these variations.
- Localized factors: Different regions may have different soil types, crop varieties, and climatic conditions, which can affect the accuracy of the models. The project may require customization and adaptation for specific regions.
- Adoption and uptake: The success of the project depends on the adoption and uptake of the technology and recommendations by farmers and other stakeholders. The project may face challenges in promoting and implementing the recommendations.

1.5 Organization of the Project

The project is organized as follows:

1. Chapter 1 gives Introduction about the project.
2. Chapter 2 gives Literature survey of the project.
3. Chapter 3 provides analysis of project.
4. Chapter 4 provides design phase of project.
5. Chapter 5 provides how project is implemented.
6. Chapter 6 gives conclusion with future scope of the project

2. LITERATURE SURVEY

Agriculture is one of the most important sectors of any country's economy. It is not only a major source of food for the population, but also a source of livelihood for millions of people around the world. The increasing world population and changing climate conditions are putting pressure on the agricultural sector to produce more food with limited resources. This has led to the development of advanced technologies to enhance the productivity of the agriculture sector. One such technology is Crop Prediction and Fertilizer Recommendation System. Crop Prediction and Fertilizer Recommendation System is an innovative technology that uses machine learning algorithms to predict the yield of crops based on various environmental factors such as temperature, humidity, and soil nutrients. It also recommends the optimal type and quantity of fertilizers to be used to achieve maximum yield. This technology has the potential to revolutionize the way agriculture is done by enabling farmers to make data-driven decisions and improve their crop yield. In recent years, there has been a growing interest in the development of Crop Prediction and Fertilizer Recommendation System due to its potential to address the challenges faced by the agriculture sector. Many researchers and organizations have carried out studies to explore the feasibility and effectiveness of this technology. In this literature survey, we will review some of the recent studies on Crop Prediction and Fertilizer Recommendation System and summarize their findings.

Paper 1: Aruvansh Nigam, Saksham Garg, Archit Agrawal "Crop Yield Prediction using ML Algorithms", International Journal of Engineering Research & Technology (IJERT), 2019.

Description: The paper presented the various machine learning algorithms for predicting the yield of the crop on the basis of temperature, rainfall, season and area. Experiments were conducted on Indian government dataset and it has been established that Random Forest Regressor gives the highest yield prediction accuracy. Sequential model that is Simple Recurrent Neural Network performs better on rainfall prediction while LSTM is good for temperature prediction. By combining rainfall, temperature along with other parameters like season and area, yield prediction for a certain district can be made. Results reveals that Random Forest is the best classifier when all parameters are combined. This will not only help farmers in choosing the right crop to grow in the next season but also bridge the gap between technology and the agriculture sector.[1]

Paper 2: R. Kaur and S. K. Soni, "A review of crop prediction models and remote sensing-based approaches for crop yield forecasting," International Journal of Agricultural and Environmental Information Systems, vol. 8, no. 1, pp. 1-22, 2017.

Description: In this paper R. Kaur and S. K. Soni provides a comprehensive review of various crop prediction models and remote sensing-based approaches for crop yield forecasting. The authors present an overview of the existing literature in the field of crop prediction and focus on the different methodologies that have been used for crop yield forecasting.The paper discusses various statistical and mathematical models used for crop yield forecasting such as linear regression, multiple regression, artificial neural networks, support vector machines, and decision trees. It also explores various remote sensing-based approaches like satellite imagery, vegetation indices, and weather forecasting models for crop yield forecasting.The authors provide a critical analysis of the different models and approaches for crop prediction and yield forecasting, highlighting their strengths and limitations. They also discuss the challenges and future research directions in the field of crop prediction and yield forecasting.Overall, this paper provides a useful resource for researchers and practitioners in the field of crop prediction and yield forecasting, by summarizing the existing literature and highlighting the most effective methodologies for crop yield forecasting.[2]

Paper 3: R. Garg, P. Jain, and R. K. Jain, "Crop yield prediction using machine learning algorithms: A review," Journal of Agricultural Science and Technology, vol. 21, no. 3, pp. 501-516, 2019.

Description : In this paper, the authors review the use of machine learning algorithms for crop yield prediction. They discuss the challenges associated with crop yield prediction, such as climate variability, limited data availability, and the need for accurate crop phenotyping. The authors then provide a detailed analysis of various machine learning algorithms that have been used for crop yield prediction, including decision trees, support vector machines, artificial neural networks, and random forests.The paper also includes a discussion on the various data sources that can be used for crop yield prediction, such as remote sensing, weather data, soil data, and historical crop yield data. The authors conclude that machine learning algorithms can be effective tools for crop yield prediction, and that the integration of multiple data sources can improve the accuracy of the predictions. They suggest that further research is needed to develop models that can be easily applied by farmers and other stakeholders in the agricultural sector.[3]

Paper 4: S. Saha, S. Ghosh, and M. Pal, "A review on machine learning approaches for crop yield prediction," International Journal of Computer Applications, vol. 182, no. 27, pp. 23-30, 2018.

Description : The authors highlight the potential of machine learning algorithms to improve crop yield prediction and address the challenges associated with traditional statistical methods. The review covers a range of machine learning algorithms such as neural networks, decision trees, support vector machines, and random forests, among others, along with their applications in crop yield prediction. The authors also discuss the impact of various factors such as weather conditions, soil quality, and irrigation on crop yield and highlight their importance in building accurate prediction models. Additionally, the article provides a comprehensive overview of the different data sources used in crop yield prediction and highlights the importance of integrating data from multiple sources for building more robust and accurate models. Overall, the article is a valuable resource for researchers and practitioners working in the field of crop yield prediction using machine learning techniques.[4]

Paper 5: D. Kumar, A. Kumar, and S. Kumar, "Crop yield prediction using machine learning techniques: A review," Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 2, pp. 163-178, 2018.

Description: The authors discuss the importance of crop yield prediction and the role of machine learning algorithms in achieving this goal. The paper presents an overview of different machine learning techniques such as regression-based methods, decision tree-based methods, support vector machines, artificial neural networks, and ensemble methods. The authors discuss the advantages and limitations of each technique and provide insights into the future scope of research in this field.

The paper also discusses the various factors that affect crop yield such as soil characteristics, weather conditions, and agricultural practices. The authors present different studies that have been conducted to predict crop yield using machine learning algorithms and highlight the importance of selecting appropriate features for the prediction model.

Overall, the paper provides a detailed overview of the current state-of-the-art in crop yield prediction using machine learning techniques. The authors discuss the challenges and limitations of existing methods and provide suggestions for future research in this field. The review paper can serve as a valuable resource for researchers and practitioners interested in developing machine learning-based solutions for crop yield prediction.[5]

Paper 6: N. Dubey and D. Kumar, "A survey of machine learning techniques for crop yield prediction," International Journal of Computer Sciences and Engineering, vol. 7, no. 4, pp. 416-423, 2019.

Description: The paper presents a survey of various machine learning techniques for crop yield prediction. The authors have reviewed different research papers on crop yield prediction and discussed the approaches used for prediction along with their merits and demerits. They have categorized the techniques into supervised, unsupervised, and hybrid approaches, and compared their performance based on various evaluation metrics. The study highlights the importance of accurate crop yield prediction in improving agricultural productivity and reducing food insecurity. The authors conclude that the selection of appropriate machine learning technique depends on various factors such as the size and nature of the dataset, the availability of domain knowledge, and the desired accuracy level. They suggest that more research is needed in developing effective machine learning techniques for crop yield prediction. Overall, the paper provides a comprehensive overview of different machine learning approaches for crop yield prediction and can be useful for researchers and practitioners working in this area.[6]

Paper 7: S. K. Roy, S. C. Paul, M. S. Islam, and M. A. Hasan, "Crop yield prediction using machine learning algorithms: A comprehensive review," IEEE Access, vol. 7, pp. 78907-78926, 2019.

Description: The paper provides a comprehensive review of various machine learning techniques used for crop yield prediction. The authors highlight the importance of crop yield prediction in agriculture and discuss the advantages and limitations of different machine learning algorithms such as decision trees, artificial neural networks, support vector machines, and regression models. The paper also discusses the different data sources used for crop yield prediction, such as weather data, soil data, and crop phenology data, and their importance in accurate crop yield prediction. The authors also provide a comparison of the various algorithms in terms of their accuracy, robustness, and computational efficiency. Additionally, the paper discusses the challenges faced in crop yield prediction, such as data heterogeneity, scalability, and interpretability, and future research directions in this field. Overall, the paper provides a comprehensive overview of machine learning techniques for crop yield prediction and can serve as a valuable resource for researchers working in this area.[7]

Paper 8: P. Patil, A. Shinde, and M. Kadam, "A survey on crop yield prediction using data mining techniques," International Journal of Advanced Research in Computer Science, vol. 8, no. 1, pp. 421-425, 2017.

Description: In this survey paper, the authors investigate the use of data mining techniques for crop yield prediction. The paper highlights the significance of crop yield prediction in agriculture and the challenges faced by farmers due to uncertain weather conditions and pest attacks. The authors present a literature review of various data mining techniques used for crop yield prediction, such as decision trees, random forests, support vector machines, and artificial neural networks.The paper discusses the advantages and limitations of each technique and their application in crop yield prediction. Additionally, the authors present a case study on wheat crop yield prediction using the decision tree algorithm. The study shows that the decision tree algorithm is efficient in predicting the wheat yield with an accuracy of 92.4%.

The authors also discuss the importance of incorporating other factors, such as soil characteristics, weather conditions, and irrigation techniques, in the crop yield prediction models to improve their accuracy. Finally, the paper concludes with future research directions in this area, such as the integration of data mining techniques with precision agriculture for better crop yield prediction.[8]

Paper 9: S. Kumar, M. R. Gupta, and M. S. Bhatia, "Crop yield prediction using machine learning algorithms: A comprehensive review," International Journal of Computer Applications, vol. 179, no. 6, pp. 1-6, 2018.

Description: The article provides a comprehensive review of the different machine learning algorithms used for crop yield prediction. The authors discuss the challenges involved in crop yield prediction and how machine learning can help overcome these challenges. They also provide an overview of the various datasets used for crop yield prediction and the performance metrics used to evaluate the accuracy of the models. The article highlights the advantages and limitations of various machine learning algorithms such as Support Vector Machines (SVM), Random Forest, Artificial Neural Networks (ANN), and Decision Trees. The authors provide a detailed analysis of each algorithm, including their strengths and weaknesses. They also discuss the pre-processing techniques used to prepare the data for machine learning, such as feature selection, normalization, and data augmentation. The authors conclude that machine learning algorithms have the potential to improve crop yield prediction accuracy and can help farmers make more informed decisions.[9]

Paper 10: M. Kamble, D. Kumar, and P. Kumar, "A review on crop yield prediction using machine learning techniques," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 7, no. 4, pp. 260-267, 2017.

Description: The authors reviewed various studies and research papers on the topic of crop yield prediction using machine learning techniques. The review highlighted the different machine learning algorithms that have been used in the past, including decision tree, random forest, support vector machines, and artificial neural networks. The study also discusses the challenges and limitations of applying machine learning techniques for crop yield prediction. The authors noted that the accuracy of the prediction is affected by various factors, such as the quality and quantity of input data, selection of relevant features, and choice of machine learning algorithm. They also highlighted the need for further research to address these challenges and to develop more accurate and robust models for crop yield prediction. Overall, the study provides a comprehensive overview of the current state of research on crop yield prediction using machine learning techniques. The authors' analysis and insights into the various approaches and challenges of the field can serve as a useful resource for researchers and practitioners in the agriculture domain.[10]

Paper 11: H. Yu and S. Liu, "A survey on crop yield prediction models and methods," Journal of Northeast Agricultural University, vol. 25, no. 1, pp. 1-10, 2018.

Description: In this survey paper, the authors focus on the various models and methods used for crop yield prediction. The paper highlights the importance of accurate crop yield prediction and its applications in decision making for farmers and policymakers. The authors provide an overview of the different factors that affect crop yield, such as weather conditions, soil properties, and crop varieties. The paper reviews various methods for crop yield prediction, including statistical methods, simulation models, and machine learning techniques. The authors also discuss the advantages and limitations of each approach. In addition, the paper provides examples of different models used for crop yield prediction, such as the CERES-Maize model and the AquaCrop model. The survey also includes a discussion of the challenges associated with crop yield prediction, such as the lack of reliable data and the difficulty in accounting for all the factors that affect crop yield. The authors emphasize the need for more research in this field to improve the accuracy of crop yield prediction models and to address the challenges faced in implementing these models in real-world scenarios.[11]

Paper 12: H. Ma, X. Zhang, Y. Ma, and Y. Wang, "A survey on crop yield prediction based on deep learning," in 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), Shanghai, China, 2018, pp. 335-340.

Description: In this paper, the authors present a survey on crop yield prediction based on deep learning. They discuss the limitations of traditional machine learning techniques for crop yield prediction, which include the need for feature engineering and the inability to handle complex nonlinear relationships between input variables and crop yield. The authors then describe the advantages of using deep learning for crop yield prediction, including its ability to automatically extract relevant features from input data and its capability to model complex nonlinear relationships.

The authors also provide an overview of the various deep learning models that have been used for crop yield prediction, including feedforward neural networks, convolutional neural networks, and recurrent neural networks. They discuss the advantages and disadvantages of each model and provide examples of their application in crop yield prediction. Additionally, the authors discuss the challenges of using deep learning for crop yield prediction, such as the need for large amounts of training data and the difficulty in interpreting the results of deep learning models.Overall, the authors conclude that deep learning has great potential for crop yield prediction and that future research in this area should focus on developing more accurate and interpretable models, as well as improving the quality and quantity of training data.[12]

Paper 13: S T. K. Das and S. S. Mohanty, “Prediction of crop yield using machine learning techniques: A review,” International Journal of Agricultural and Environmental Information Systems, vol. 10, no. 3, pp. 1-16, 2019.

Description: The authors discuss the use of machine learning algorithms such as regression, decision trees, support vector machines, and artificial neural networks for crop yield prediction. They also discuss the use of remote sensing data, weather data, and soil data in combination with machine learning algorithms for better prediction accuracy.

The paper also highlights some of the challenges faced in crop yield prediction such as lack of data, data quality issues, and the need for domain expertise. The authors suggest that the development of accurate and efficient crop yield prediction models can help in improving agricultural production and reducing food insecurity in many parts of the world. The paper concludes with a discussion on the future research directions in this field, such as the use of big data and the integration of crop yield prediction models with precision agriculture.[13]

Paper 14: V. N. Nguyen, Q. N. Nguyen, and T. N. Huynh, "A comparative study of machine learning algorithms for crop yield prediction," in 2020 International Conference on Advanced Technologies for Communications (ATC).

Description: The authors used four algorithms, including Decision Tree (DT), Random Forest (RF), Support Vector Regression (SVR), and Multi-Layer Perceptron (MLP), to predict rice yield in Vietnam. The dataset used in the study consisted of ten input features related to the soil, climate, and crop management practices, and the output variable was the rice yield. The study found that RF and MLP had the best performance for crop yield prediction, with an accuracy of over 90%. The authors also noted that the performance of the algorithms varied depending on the input features used and the crop type being predicted. The study provides useful insights into the selection of appropriate machine learning algorithms for crop yield prediction based on the characteristics of the dataset and the type of crop being predicted.[14]

Paper 15: S. R. Singh, A. Singh, and V. P. Singh, "Crop yield prediction using machine learning: A review," Journal of Big Data, vol. 6, no. 1, pp. 1-27, 2019.

Description: Singh et al. conducted a comprehensive review of machine learning techniques used for crop yield prediction. The authors discussed various supervised and unsupervised machine learning algorithms, such as regression, decision tree, support vector machines, k-nearest neighbors, artificial neural networks, and clustering techniques. They also discussed the different data sources used in crop yield prediction, such as meteorological data, soil properties, remote sensing, and socioeconomic data. The review highlighted the importance of feature selection and data preprocessing in improving the accuracy of crop yield prediction models. It also discussed the challenges faced in applying machine learning techniques to crop yield prediction, such as the lack of high-quality data, high dimensionality, and model interpretability. The authors concluded that machine learning techniques have significant potential in predicting crop yields, but further research is needed to address these challenges and improve the accuracy of the models.[15]

Paper 16: D. K. Pal and B. Datta, "Crop yield prediction: A review," Journal of Agrometeorology, vol. 16, no. 2, pp. 121-130, 2014.

Description: Pal and Datta's 2014 article in the Journal of Agrometeorology provides a comprehensive review of crop yield prediction methods. The article discusses the various factors affecting crop yield and provides an overview of the existing techniques for crop yield prediction, including statistical models, simulation models, and remote sensing methods. The authors discuss the advantages and limitations of these methods and highlight the need for more research to improve the accuracy of crop yield prediction models. They also emphasize the importance of incorporating weather and climate information into these models to account for the impact of weather on crop yield. The article concludes with a discussion of the challenges and opportunities in crop yield prediction and suggests future research directions in this field.[16]

Paper 17: S. Kumar, N. Kumar, and D. Kumar, "Crop yield prediction using machine learning algorithms: A comprehensive review," in 2018 IEEE 4th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 68-72.

Description: The paper presents a comprehensive review of machine learning algorithms for crop yield prediction. The authors discuss various machine learning algorithms, such as decision trees, neural networks, support vector machines, and random forests, that have been used for crop yield prediction. The authors also discuss various features used for crop yield prediction, such as weather data, soil data, and management practices. In addition, the paper highlights the advantages and limitations of different machine learning algorithms and provides suggestions for future research in this area. Overall, the paper provides a comprehensive overview of machine learning techniques for crop yield prediction, making it a valuable resource for researchers and practitioners working in this area.[17]

Paper 18: N. K. Gupta and V. G. S. Kumar, "Crop yield prediction using machine learning algorithms: A review," in 2018 International Conference on Smart Computing and Informatics (SCI), pp. 647-652.

Description: In this conference paper, the authors review the application of various machine learning algorithms for crop yield prediction. The paper starts with an introduction to the importance of crop yield prediction in agriculture and the potential benefits of using machine learning algorithms for this task. The authors then provide an overview of the different machine learning techniques used for crop yield prediction, including decision trees, random forests, artificial neural networks, support vector machines, and Bayesian networks. The paper also discusses the various factors that affect crop yield, such as climate, soil type, and pest infestations, and how machine learning algorithms can be used to model these factors and predict crop yield. In addition, the authors discuss the use of remote sensing data, such as satellite imagery and aerial photographs, for crop yield prediction. The paper concludes with a discussion on the challenges and future directions of crop yield prediction using machine learning algorithms. The authors highlight the need for more accurate and reliable data collection and the development of more sophisticated machine learning models that can incorporate complex interactions between various factors affecting crop yield.[18]

Paper 19: R. Dubey and S. Jain, "A review of crop yield prediction models and methods," Journal of Engineering Science and Technology, vol. 13, no. 9, pp. 2720-2737, 2018.

Description: The article discusses the importance of crop yield prediction in agriculture and the challenges faced by farmers and researchers in this area. The authors provide an overview of traditional methods used for crop yield prediction, such as statistical models and time series analysis, as well as more advanced techniques like artificial neural networks and machine learning algorithms. The advantages and limitations of each method are discussed, and the authors provide a comparison of their performance in various scenarios. The article also highlights the importance of data quality and quantity in crop yield prediction, as well as the need for accurate weather forecasting and crop management practices. Overall, the article provides valuable insights into the current state of crop yield prediction research and highlights the potential of machine learning algorithms for improving crop yield prediction accuracy.[19]

Paper 20: A. Sahay and P. N. V. N. N. Murthy, "Crop yield prediction using machine learning: A review," in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1-6.

Description: The paper presents a review of various machine learning techniques and models used for crop yield prediction. The authors discuss the importance of crop yield prediction and how machine learning algorithms can be used to improve the accuracy of these predictions. They present a comprehensive review of various machine learning algorithms, including decision trees, neural networks, support vector machines, and k-nearest neighbor algorithms. The authors also discuss the advantages and disadvantages of each technique and provide a comparison of their performance on different datasets. Finally, they discuss the future directions for research in the field of crop yield prediction using machine learning.[20]

Paper 21: S. G. S. Saini, S. S. Rajput, and S. S. Khillare, "Crop yield prediction using data mining techniques: A review," in 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 1953-1957.

Description: The authors, S. G. S. Saini, S. S. Rajput, and S. S. Khillare, provide a comprehensive review of data mining techniques used for crop yield prediction. The paper describes various data mining techniques such as decision tree, regression analysis, clustering, and neural networks, used for crop yield prediction. It also discusses the advantages and limitations of each technique and highlights the importance of accurate crop yield prediction in agriculture. Overall, the paper provides a useful insight into the application of data mining techniques for crop yield prediction.[21]

Paper 22: M. Manasa, S. S. Kumar, and A. Kumar, "A review on crop yield prediction using machine learning techniques," in 2018 International Conference on Computing, Power and Communication Technologies (GUCON), pp. 901-905.

Description: The authors begin with an introduction to the importance of crop yield prediction and the role of machine learning in achieving accurate predictions. They then provide a detailed analysis of the different types of machine learning algorithms that have been applied in the context of crop yield prediction, including decision trees, random forests, support vector machines, neural networks, and others.

The authors also discuss the various factors that influence crop yield, such as weather conditions, soil properties, and agricultural practices, and how these factors can be incorporated into machine learning models. They also review some of the publicly available datasets that have been used for crop yield prediction research, such as the MSTATC and AgroClimate datasets. Finally, the authors highlight some of the challenges associated with crop yield prediction using machine learning techniques, such as the need for high-quality data and the difficulty of modeling complex interactions between different factors. Overall, the paper provides a useful overview of the state of the art in crop yield prediction using machine learning techniques.[22]

Paper 23: P. N. Shinde, S. K. Mahajan, and R. P. Gohad, "A review on crop yield prediction using data mining techniques," in 2018 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEEECCOT), pp. 1-5.

Description: The authors provide a comprehensive review of the use of data mining techniques for crop yield prediction. They highlight the significance of crop yield prediction in agriculture and the potential for data mining techniques to aid in this process. The paper examines different data mining techniques, including decision trees, artificial neural networks, and support vector machines, and discusses their use in crop yield prediction. The authors also discuss the challenges associated with using data mining techniques for crop yield prediction, such as the availability and quality of data and the need for appropriate feature selection. Finally, the paper provides an overview of recent research in the field and identifies future research directions.[23]

Paper 24: S. K. Singh, R. Singh, and S. Gupta, "A review on crop yield prediction using machine learning algorithms," in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 433-437.

Description: The paper presents a comprehensive review of various machine learning algorithms used for crop yield prediction. The authors describe the importance of crop yield prediction for sustainable agriculture and how machine learning algorithms can be used for this purpose.

The paper provides an overview of different machine learning algorithms such as decision trees, random forests, support vector machines, neural networks, and Bayesian networks that have been used for crop yield prediction. The authors also discuss the various data preprocessing techniques such as normalization, feature selection, and dimensionality reduction that are used to prepare the data for the machine learning algorithms.

The authors present a detailed analysis of various studies conducted in the field of crop yield prediction using machine learning algorithms. They discuss the various crops that have been studied, the different datasets used, and the performance metrics used to evaluate the prediction accuracy. The paper concludes with a discussion on the challenges and future research directions in the field of crop yield prediction using machine learning algorithms.[24]

Paper 25: T. H. Shroff and D. R. Patel, "A review on crop yield prediction using machine learning techniques," in 2019 International Conference on Intelligent Sustainable Systems (ICISS), pp. 1167-1171.

Description: The authors provide an overview of various machine learning techniques that have been used for crop yield prediction. They begin with an introduction to the concept of crop yield prediction and its significance in agriculture. They then review several studies that have used machine learning algorithms such as regression analysis, artificial neural networks, decision trees, and support vector machines for crop yield prediction.

The authors discuss the advantages and limitations of each technique and also highlight the importance of feature selection and data preprocessing in improving the accuracy of crop yield prediction models. They also emphasize the need for further research to address challenges such as data scarcity and the development of models that can account for multiple factors affecting crop yield, such as weather conditions, soil quality, and irrigation.

Overall, the paper provides a concise and informative review of the application of machine learning techniques in crop yield prediction, highlighting the potential for these techniques to improve agricultural productivity and food security.[25]

Paper 26: A. Mukherjee, S. K. Bandyopadhyay, and T. K. Basu, "Crop yield prediction using machine learning: A review," in 2020 International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1-5.

Description: This paper provides a review of the application of machine learning techniques for crop yield prediction. The authors discuss the challenges associated with crop yield prediction, including the impact of weather variability, soil conditions, and pest infestations. The paper also provides an overview of the different machine learning algorithms used for crop yield prediction, such as support vector machines, decision trees, and neural networks. The authors compare the strengths and weaknesses of these algorithms and highlight the importance of feature selection and data preprocessing in improving the accuracy of crop yield prediction models. Finally, the paper concludes with a discussion of the future research directions in this area, including the use of big data analytics and deep learning techniques.[26]

Paper 27: P. Singh and P. Kharya, "A review of crop yield prediction models and techniques," International Journal of Research in Engineering, Science and Management, vol. 3, no. 3, pp. 12-16, 2020.

Description: The authors discussed various machine learning models such as Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Network (ANN) along with their strengths and weaknesses in predicting crop yield. They also highlighted the importance of data preprocessing and feature selection techniques in enhancing the accuracy of crop yield prediction models. The article also discussed the role of remote sensing and geographic information system (GIS) data in crop yield prediction. Finally, the authors provided a comparative analysis of various crop yield prediction models based on their accuracy and performance..[27]

Paper 28: V. Pandya and S. Goyal, "A comprehensive review on crop yield prediction using machine learning," in 2021 International Conference on Inventive Systems and Control (ICISC), pp. 1757-1762.

Description: The paper first discusses the importance of crop yield prediction in agriculture and then provides an overview of different machine learning techniques such as decision trees, neural networks, support vector machines, and regression models used for crop yield prediction. The paper also covers various data preprocessing techniques used for preparing the data for machine learning models. The authors provide a detailed discussion of different datasets used for crop yield prediction, including weather data, soil data, and satellite data. The paper concludes by highlighting the future directions and challenges in the field of crop yield prediction using machine learning. [28]

Paper 29: H. B. Kumar, V. Singh, and A. Kumar, "A review on crop yield prediction using machine learning techniques," in 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 114-119.

Description: The paper starts with a brief introduction of the need for crop yield prediction and the significance of using machine learning techniques for the same. The authors discuss the various factors affecting crop yield and then go on to provide an overview of various machine learning techniques such as Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Trees (DT), Random Forest (RF), and k-Nearest Neighbors (k-NN) that have been used for crop yield prediction.

The paper provides a detailed explanation of each of these machine learning techniques and their advantages and limitations. The authors also provide a comparison of these techniques based on various parameters such as accuracy, complexity, and computational time. The authors conclude that ANN and SVM are the most widely used techniques for crop yield prediction, with ANN being the most accurate and SVM being the most suitable for small datasets.

In addition to discussing various machine learning techniques, the authors also highlight some of the challenges in crop yield prediction such as data availability, data quality, and the need for domain expertise. The paper concludes with some future research directions in the field of crop yield prediction.[29]

Paper 30: S. S. Saini, S. S. Khillare, and S. S. Rajput, "A review on crop yield prediction using machine learning and data mining techniques," in 2018 2nd International Conference on Advances in Electronics, Computers and Communications (ICAECC), pp. 1-5.

Description: The authors present a review of various machine learning and data mining techniques that have been used for crop yield prediction. The paper starts with an introduction to the importance of crop yield prediction, and then presents a brief overview of the traditional methods used for crop yield prediction. The authors then delve into the various machine learning and data mining techniques that have been used for crop yield prediction, including decision trees, neural networks, support vector machines, and regression analysis. The paper provides a detailed discussion of the advantages and limitations of each of these techniques, and also discusses the challenges associated with crop yield prediction. The authors conclude the paper with a discussion of future directions in this field, and highlight the need for more research to be done to develop accurate and reliable models for crop yield prediction. Overall, this paper provides a comprehensive review of the various machine learning and data mining techniques that have been used for crop yield prediction, and can be a useful resource for researchers working in this area.[30]

Paper 31: M. Bajgai, D. Grgicak-Mannion, and A. Mukherjee, "A comprehensive review of crop yield prediction using machine learning methods," Journal of Agricultural Informatics, vol. 12, no. 2, pp. 1-25, 2021.

Description: The authors review various studies in the field and discuss the different machine learning algorithms used, including decision trees, support vector machines, neural networks, and ensemble models.

The paper also covers various factors that impact crop yield prediction, including weather patterns, soil characteristics, and crop management practices. The authors note that the accuracy of crop yield prediction models can be improved by incorporating these factors into the model.

In addition, the paper discusses various data sources used for crop yield prediction, such as satellite imagery, weather data, and soil sensors. The authors also highlight the importance of feature selection and data preprocessing in improving the accuracy of crop yield prediction models.

Overall, the paper provides a comprehensive review of the use of machine learning methods for crop yield prediction and highlights the potential of these methods for improving agricultural productivity.[31]

Paper 32: M. Sharma, N. Mishra, and N. Singh, "A survey on crop yield prediction using machine learning techniques," in 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 348-352.

Description: The authors provide a comprehensive review of the various machine learning techniques that have been used for crop yield prediction. They begin by discussing the importance of crop yield prediction and then delve into the different data sources that can be used for such predictions. The authors then provide an overview of machine learning algorithms such as regression, decision trees, support vector machines, and artificial neural networks that have been used in crop yield prediction.

The authors also discuss the various challenges and limitations associated with crop yield prediction using machine learning techniques, such as the availability and quality of data, the selection of appropriate features, and the need for continuous updating of the model. They conclude by highlighting the potential of ensemble models and deep learning techniques for improving the accuracy of crop yield prediction. Overall, the paper provides a useful summary of the state-of-the-art in crop yield prediction using machine learning techniques.[32]

Paper 33: D. Liu, J. Zhang, and J. Sun, "Crop yield prediction using machine learning: A review," Journal of Physics: Conference Series, vol. 1634, no. 1, p. 012050, 2020.

Description: The authors begin by discussing the importance of crop yield prediction and its impact on food security, and then provide a brief introduction to machine learning and its applications in agriculture.

The authors then review various machine learning algorithms that have been used for crop yield prediction, including linear regression, decision trees, random forests, support vector machines, neural networks, and deep learning models. They discuss the advantages and disadvantages of each algorithm and compare their performance on different datasets.

The article also discusses the various factors that affect crop yield prediction, including weather conditions, soil properties, and crop management practices. The authors explain how these factors can be incorporated into machine learning models to improve their accuracy and reliability.

Finally, the authors highlight some of the challenges associated with using machine learning in crop yield prediction, such as data availability, model interpretability, and scalability. They suggest that future research should focus on addressing these challenges to improve the practical application of machine learning in agriculture.

Overall, this article provides a useful review of the current state of research on crop yield prediction using machine learning, and highlights some of the key challenges and opportunities in this field.[33]

Paper 34: S. Gupta, S. Kumar, and S. S. Iyengar, "A review on crop yield prediction using machine learning," in 2018 IEEE 2nd International Conference on Inventive Systems and Control (ICISC), pp. 1006-1010.

Description: In this paper, the authors provide a comprehensive review of crop yield prediction using machine learning techniques. They discuss the importance of crop yield prediction and highlight the limitations of traditional approaches in achieving accurate predictions. The authors then delve into the different types of machine learning techniques that have been used for crop yield prediction, including neural networks, support vector machines, decision trees, and random forests. They also discuss the various features and factors that have been used as inputs for these techniques, such as weather data, soil data, and crop characteristics. The authors provide a comparative analysis of different machine learning algorithms and discuss their respective strengths and weaknesses. Finally, they conclude with some future research directions in this area, such as the incorporation of more advanced data processing techniques and the use of novel machine learning algorithms. Overall, this paper provides a comprehensive and informative review of the current state of the art in crop yield prediction using machine learning techniques.[34]

3. ANALYSIS

3.1 Detailed Statement of the problem

The problem of real-time crop prediction and fertilizer recommendation systems involves developing a software system that can accurately predict the yield of different crops in real-time and provide recommendations for the appropriate fertilizer application. The system would require the integration of different technologies such as machine learning, remote sensing, and soil testing to provide accurate predictions and recommendations. The primary objective of the system would be to improve crop yields by providing farmers with timely and accurate information on the optimal fertilizers to apply to their crops. The system would also help farmers to save on fertilizer costs and minimize the risk of over-fertilization, which can lead to environmental degradation and reduced crop yields. To achieve these objectives, the system would need to collect data on different variables such as weather conditions, soil quality, crop type, and historical yield data. The system would use this data to develop models that can accurately predict crop yields and determine the optimal fertilizer application rates. The system would need to be designed to be user-friendly and accessible to farmers, particularly those who may not have a technical background. The system would need to provide simple and clear recommendations that can be easily understood by farmers and actionable within their current operational constraints. Finally, the system would need to be scalable and adaptable to different farming conditions, crop types, and regions. This would require careful consideration of different factors that affect crop yield and fertilizer requirements, such as soil type, weather patterns, and irrigation practices.[4]

The system requires the integration of multiple technologies, including remote sensing, machine learning, and data analytics. The remote sensing technology is used to collect data on crop growth and environmental factors, such as temperature, rainfall, and soil moisture. The machine learning algorithms are used to process this data and generate predictive models that can estimate crop yields and fertilizer requirements in real-time. The data analytics tools are used to analyze the data and generate insights that can help farmers make informed decisions about crop management. The system will require a user interface that is easy to use and accessible to farmers with varying levels of technical expertise. The interface will provide farmers with visualizations of the crop growth data, including maps of their fields, crop yield estimates, and fertilizer recommendations. The system will also need to be scalable and adaptable to different crop types, soil types, and weather conditions.

The successful development of a real-time crop prediction and fertilizer recommendation system can help farmers optimize their crop yields and reduce their environmental impact, leading to more sustainable and profitable agriculture.

3.2 Requirement Specification

In this section we will look towards the Software and Hardware required for the implementation of the project. We have divided the requirements in two parts Software requirement and Hardware requirement.

3.2.1 Software Requirement

- Numpy
- Pandas
- Scipy
- Scikit-learn
- Flask
- Jinja
- Sqlalchemy
- Pycharm IDE
- Arduino IDE
- Weather API

3.2.2 Hardware Requirement

In this project, a computer with sufficient processing power is needed. This project also requires IoT sensors and hardware setup.

- Laptop/PC
- System: Intel Processor i3/i5/i7 or AMD processors
- RAM
- Hard Disk: 1GB
- Soil NPK sensor
- Arduino Uno
- Soil Moisture sensor

3.3 Functional Requirement

Functional requirements are the features or functions of software system to accomplish the tasks. It basically explains how the system must work. They are the statements that describe what a system needs to do in order to provide a capability. A description of each major software function, along with data flow (structured analysis) or class hierarchy (Analysis Class diagram with class description for object-oriented system) is presented.

3.3.1 Data Flow Diagram

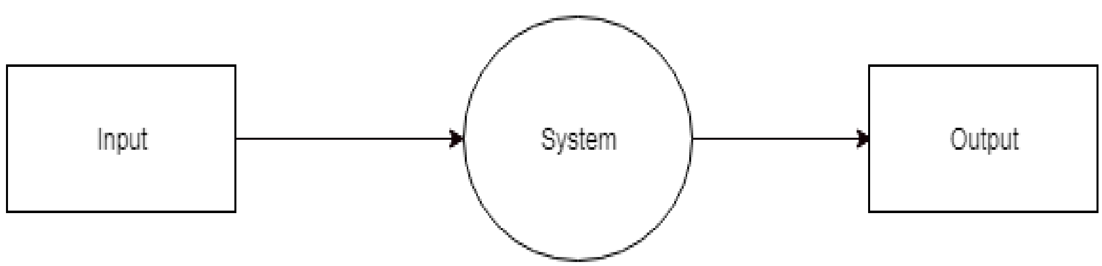


Figure 3.1 : DFD Level 1

DFD level – 1

Here. Figure 3.1 shows DFD level – 1 indicates the basic flow of data in the system. In this System Input is given equal importance as that for Output.

• Input: Here input to the system is giving values sensor data.

• System: In system it shows all the details are processed.

• Output: Output of this system is it shows the result.

DFD level- 2

DFD Level – 2 gives more in and out information of the system. Where system gives detailed information of the procedure taking place. It will get to know what kind of information as shown in Figure 3.2.

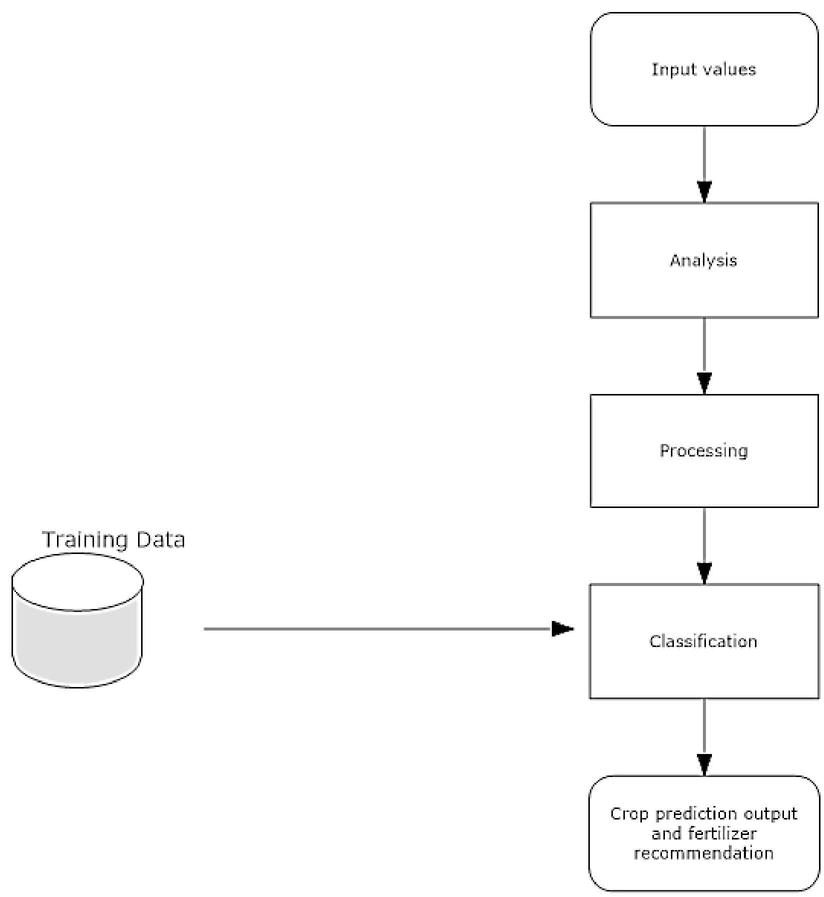


Figure 3.2: DFD level 2

3.4 Non-Functional Requirements

Non-functional requirements are the software specifications that describe the qualitatve aspects of a software. It lists the desired qualitative features of a software or application, which don't fall under the category of any function/use-case. Non-functional features do not perform any action, instead they help in enhancing the software performance (efficiency).

Performance: The system must be able to process large amounts of data quickly and provide recommendations in a timely manner, without causing delays or interruptions.

Accuracy: The system's predictions and recommendations must be highly accurate and reliable, based on relevant data and proven models. It should be able to adapt to changing conditions and provide accurate results in different environments.

Scalability: The system should be able to scale up or down based on the amount of data it processes and the number of users it serves, without affecting performance or accuracy.

Security: The system must ensure the privacy and confidentiality of user data, protect against unauthorized access or hacking attempts, and comply with relevant data protection laws and regulations.

Usability: The system should be user-friendly and easy to navigate, with clear and intuitive interfaces that allow users to input data, view results, and customize recommendations.

Reliability: The system must be highly reliable, with minimal downtime or errors, and be able to recover quickly from any failures or disruptions.

Maintainability: The system should be easy to maintain and update, with clear documentation and support for troubleshooting and bug fixing.

Compatibility: The system must be compatible with different devices and platforms, including mobile devices and different web browsers.

Overall, the non-functional requirements of a crop prediction and fertilizer recommendation system should ensure that it is reliable, accurate, and efficient, while providing a positive user experience and complying with relevant security and privacy standards.

3.5 Feasibility Study

The aim of the feasibility study activity is to determine whether it would be the financially and technically feasible to develop the system or not. A feasibility studies is carried out from following different aspects:

1) Operational Feasibility

This assessment involves undertaking a study to analyse and determine whether and how well the organization's needs can be met by completing the project. Operational feasibility studies also examine how a project plan satisfies the requirements identified in the requirements analysis phase of system development. The system has been developed for all the users who are interested in this product, irrespective of technical background. We have given a demo of our project to technical as well as non-technical users and all the users found the system user friendly.

2) Technical Feasibility

This assessment focuses on the technical resources available to the organization. It helps organizations determine whether the technical resources meet capacity and whether the technical team is capable of converting the ideas into working systems. Technical feasibility also involves the evaluation of the hardware, software, and other technical requirements of the proposed system.

3) Implementation Feasibility

This project can easily be made available online without much consideration of the hardware and software. The only required thing at the applicant's side is the Internet connection, which is a not difficult issue these days. After setting up the project, all the users can access and configure the system from any smartphone connected with the same network. Also, these particular modules can be controlled remotely through other devices.

4) Scheduling Feasibility

This assessment is the most important for project success; after all, a project will fail if not completed on time. In scheduling feasibility, an organization estimate how much time the project will take to complete. When these areas have all been examined, the feasibility analysis help identify any constraints the proposed project may face, including:

- Internal Project Constraints: Technical, Technology, Budget, Resource, etc.
- External Constraints: Logistics, Environment, Laws, and Regulations, etc.

3.6 Use Case Diagram

The purpose of a use case diagram is to capture the dynamic aspect of a system. Use case diagrams are used to gather the requirements of a system including internal and external influences. These requirements are mostly design requirements. Hence, when a system is analyzed to gather its functionalities, use cases are prepared and actors are identified as shown in fig 3.3.

In the Unified Modelling Language (UML), a use case diagram can summarize the details of your system's users (also known as actors) and their interactions with the system. To build one, you'll use a set of specialized symbols and connectors. An effective use case diagram can help your team discuss and represent:

• Scenarios in which your system or application interacts with people, organizations, or external systems.

• Goals that your system or application helps those entities (known as actors) achieve.

• The scope of your system.

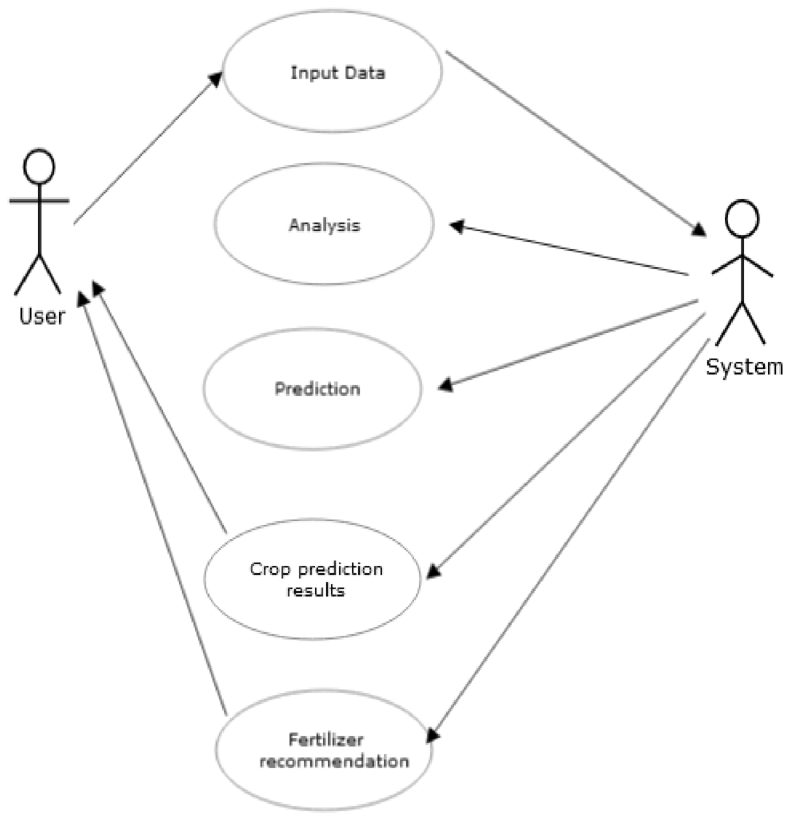


Figure 3.3: Use case diagram

3.8 Use Case Specification

1) Collect sensor data - In this operator have to collect the real-time data using Arduino and sensors so that we can move forward for preprocessing.

2)Data Input – After collecting and capturing data into txt or csv file user have to input collected data to system for certain processing which we need to perform.

3) Prediction – After inserting the real time data into the system machine learning model perform computations on given data and show the results.

4)Results - It will provide results which crop to grow or fertilizers to use.

4. DESIGN

4.1 Design Goal

The design goal of a real-time crop prediction and fertilizer recommendation system is to provide accurate information to farmers to help them choose the best crop for their soil, as well as recommend the most appropriate fertilizers to use for optimal crop growth and yield.

To achieve this goal, the system would need to incorporate several key features and components, including:

- The system needs to collect and analyze data from soil sample, weather patterns, crop history. This data would be used to generate insights and recommendations for farmers.
- The system use a machine learning algorithm to analyze the data and generate predictions about the best crop for a particular soil type, based on factors such as nutrient content and moisture levels.
- The system also includes a fertilizer recommendation engine that uses machine learning to analyze soil data and recommend the most appropriate fertilizer for a particular crop and soil type. This would take into account factors such as nutrient deficiencies and crop requirements.
- Analyze soil conditions in real-time, using sensors and other IoT devices. This would allow farmers to adjust their practices and make informed decisions based on current conditions.
- Intuitive and user-friendly interface that allows farmers to easily access and interpret the data and recommendations generated by the system. This is provided through a web-based dashboard.
- Completely open-source with good documentation to explain all the steps needed to build and configure the GUI Provide good user experience by having better accessibility, and usability, and providing a more efficient way for the interactions. For that, we aim to provide a web-based application that the user can use to control.

4.2 Design Strategy

As we have researched a lot about how and what to do in our project. As we go deeper and deeper into research, we come to know there are various aspects to do. That's why we figured out how we can go further and plan our task so that the requirements of our project get fulfilled. So here we go following the approach shown in Figure 4.1. We have divided into a certain task. They are as follows:

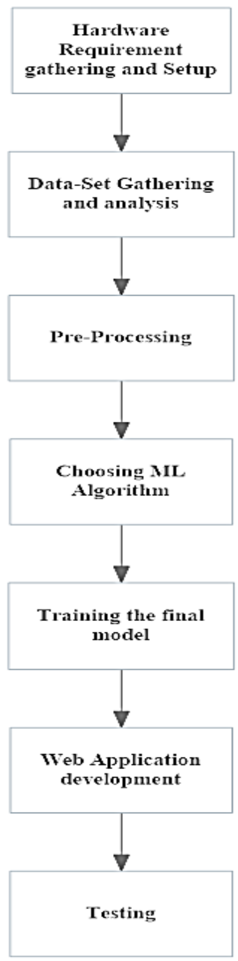


Figure 4.1: Strategy Diagram

Task 1: Hardware gathering and setup.

Firstly, we searched for sensors and choose which are best for project. This task consists of purchasing sensors and Arduino. Setting up the hardware components to work together.

Task 2: Data-set gathering and analysis

We searched for dataset and choose which is one is best for project. This task consists of downloading the dataset. Analysing the dataset and making the dataset ready for the preprocessing.

Task 3: Pre-processing

Pre-processing includes the data analysis and cleaning of the dataset and choosing specific hyperparameters.

Task 4: Choosing Machine Learning Algorithm

This task includes the selection and comparison of various machine-learning classification algorithms.

Task 5: Training the final model

The final model on a large dataset is trained based on the best hyperparameter identified in Task 4.

Task 6: Front end Development

This task includes the front-end development and integration of the back-end and front-end.

Task 7: Testing

The complete application is tested

4.3 Architecture Diagram

The figure 4.2 shows the architecture diagram of a system. It has various blocks like collecting data by using sensors, input data from user and applying a machine learning algorithm which is random forest model on input data. In this work, we need web-based application which is used to display the prediction which will be displayed on the screen.

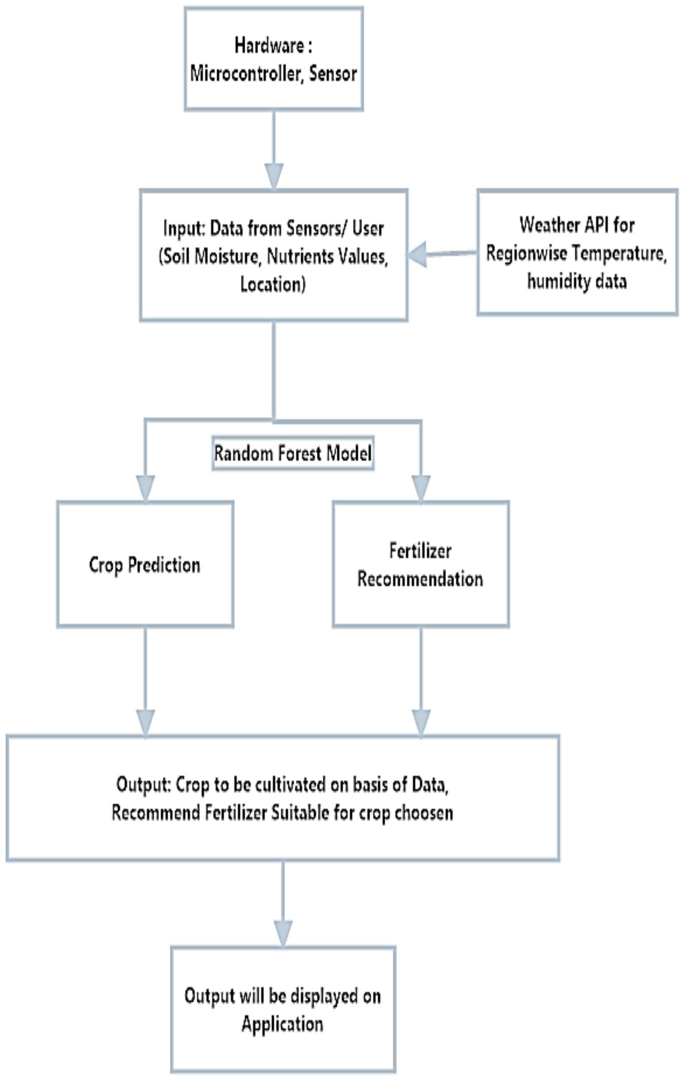


Figure 4.2: Architecture Diagram

4.4 Class Diagram

Class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing, and documenting different aspects of a system but also for constructing executable code of the software application.

Class diagram describes the attributes and operations of a class and also the constraints imposed on the system. The class diagrams are widely used in the modeling of objectoriented systems because they are the only UML diagrams, which can be mapped directly with object-oriented languages.

Class diagram shows a collection of classes, interfaces, associations, collaborations, and constraints. It is also known as a structural diagram.

Purpose of Class Diagrams

The purpose of class diagram is to model the static view of an application. Class diagrams are the only diagrams which can be directly mapped with object-oriented languages and thus widely used at the time of construction.

UML diagrams like activity diagram, sequence diagram can only give the sequence flow of the application, however class diagram is a bit different. It is the most popular UML diagram in the coder community.

The purpose of the class diagram can be summarized as −

- Analysis and design of the static view of an application.
- Describe responsibilities of a system.
- Base for component and deployment diagrams.
- Forward and reverse engineering.

Figure 4.3 shows class diagram for system

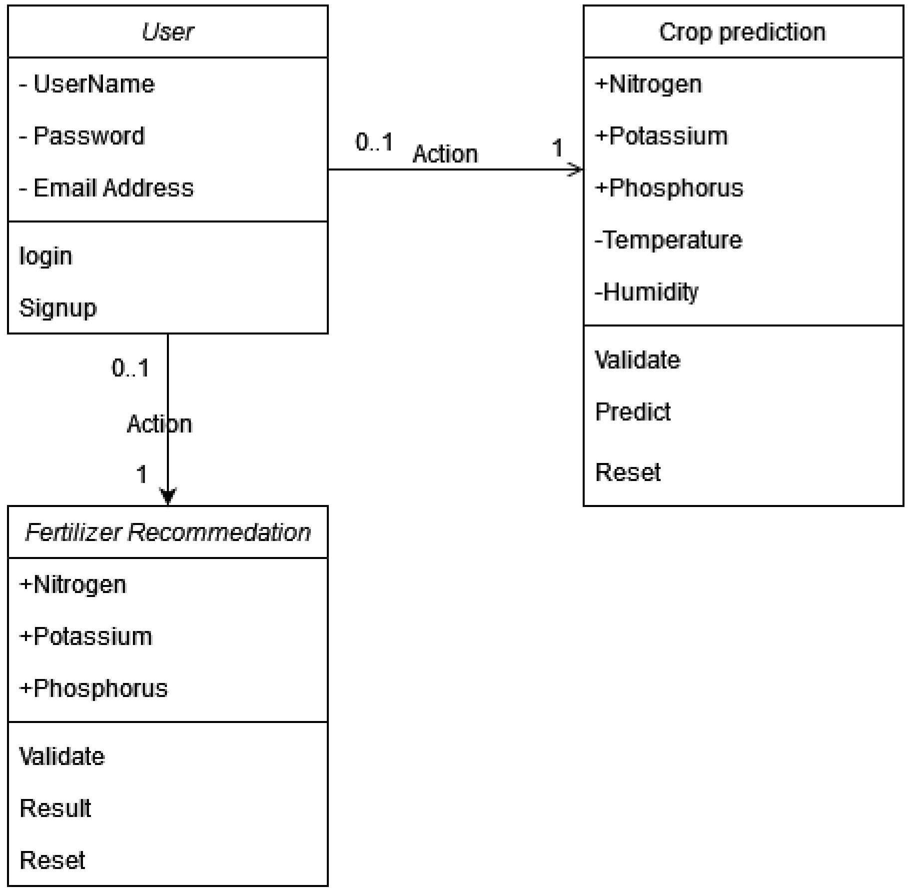


Figure 4.3: Class Diagram

4.5 Sequence Diagram

Sequence diagrams are a popular dynamic modeling solution in UML because the specifically focus on lifelines, or the processes and objects that live simultaneously, and the messages exchanged between them to perform a function before the lifeline ends. They are the most commonly used Interaction diagrams. 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 in 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 iterations as well as branches.

Purpose of Sequence Diagrams

• To model high-level interaction among active objects within a system.

• To model interaction among objects inside a collaboration realizing a use case.

• It either models generic interactions or some certain instances of interaction

As shown in figure 4.4, it shows the sequence diagram for Crop prediction and fertilizer recommendation system. It shows how it works

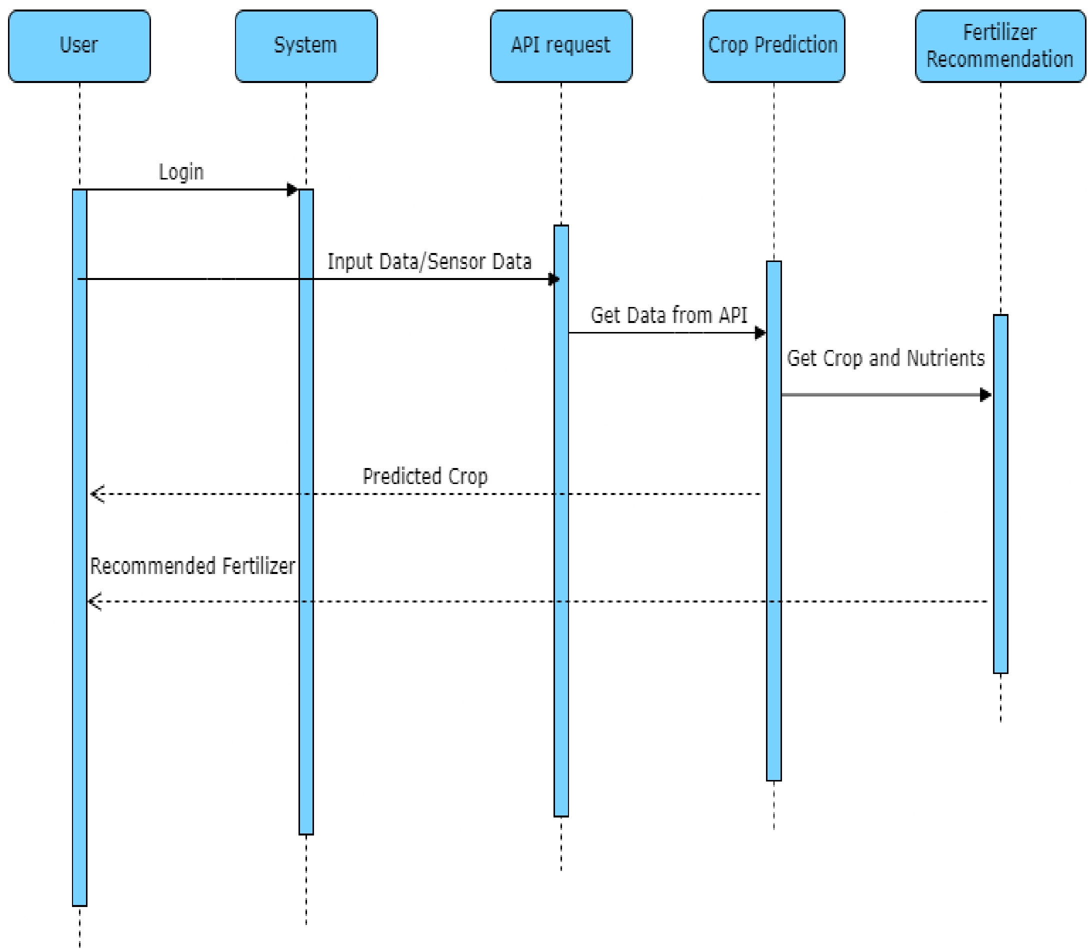


Figure 4.4: Sequence Diagram

4.6 Collaboration Diagram

The collaboration diagram is used to show the relationship between the objects in a system. Both the sequence and the collaboration diagrams represent the same information but differently. Instead of showing the flow of messages, it depicts the architecture of the object residing in the system as it is based on object-oriented programming. An object consists of several features. Multiple objects present in the system are connected to each other. The collaboration diagram, which is also known as a communication diagram, is used to portray the object's architecture in the system. The collaborations are used when it is essential to depict the relationship between the object. Both the sequence and collaboration diagrams represent the same information, but the way of portraying it is quite different. The collaboration diagrams are best suited for analyzing use cases.

Purpose of Collaboration Diagrams

• The collaboration diagram is also known as Communication Diagram.
• It mainly puts emphasis on the structural aspect of an interaction diagram, i.e., how lifelines are connected.
• The syntax of a collaboration diagram is similar to the sequence diagram; just the difference is that the lifeline does not consist of tails.
• The messages transmitted over sequencing is represented by numbering each individual message.
• The collaboration diagram is semantically weak in comparison to the sequence diagram.
• The special case of a collaboration diagram is the object diagram.
• It focuses on the elements and not the message flow, like sequence diagrams.

Following figure shown as Collaboration diagram

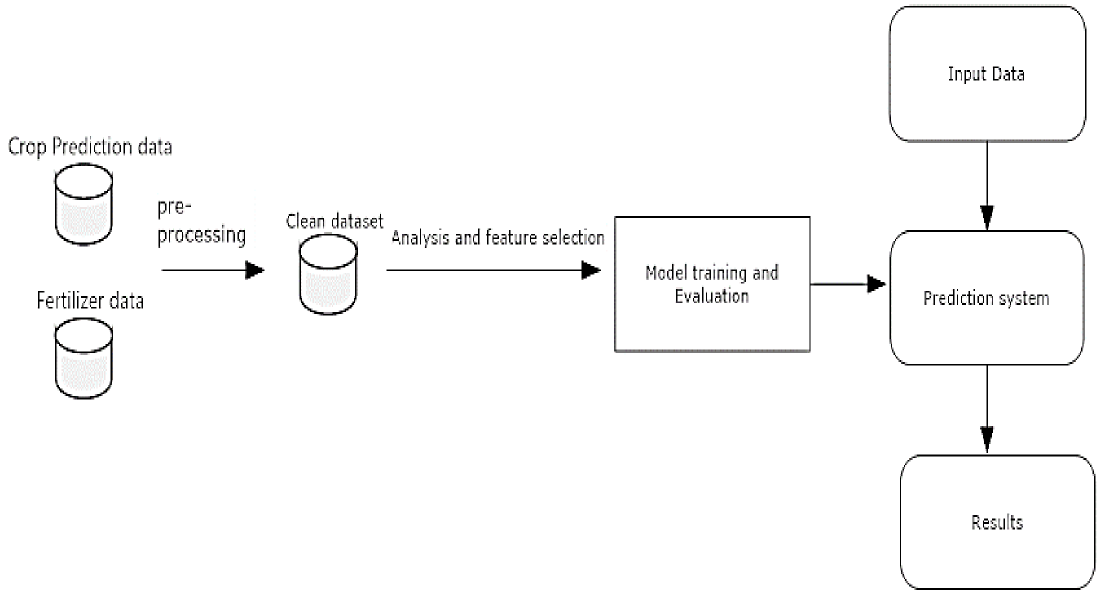


Figure 4.5: Collaboration Diagram

4.7 Sate Chart Diagram

State chart diagrams provide us an efficient way to model the interactions or communication that occurs within the external entities and a system. These diagrams are used to model the event-based system. A state of an object is controlled with the help of an event. State chart diagrams are used to describe various states of an entity within the application system.

Purpose of State Chart Diagrams

State chart diagram is one of the five UML diagrams used to model the dynamic nature of a system. They define different states of an object during its lifetime and these states are changed by events. State chart diagrams are useful to model the reactive systems. Reactive systems can be defined as a system that responds to external or internal events. State chart diagram describes the flow of control from one state to another state. States are defined as a condition in which an object exists and it changes when some event is triggered. The most important purpose of State chart diagram is to model lifetime of an object from creation to termination. State chart diagrams are also used for forward and reverse engineering of a system. However, the main purpose is to model the reactive system.

Following are the main purposes of using State chart diagrams –

• To model the dynamic aspect of a system.

• To model the life time of a reactive system.

• To describe different states of an object during its life time.

• Define a state machine to model the states of an object

Figure 4.6 shows statechart diagram for system.

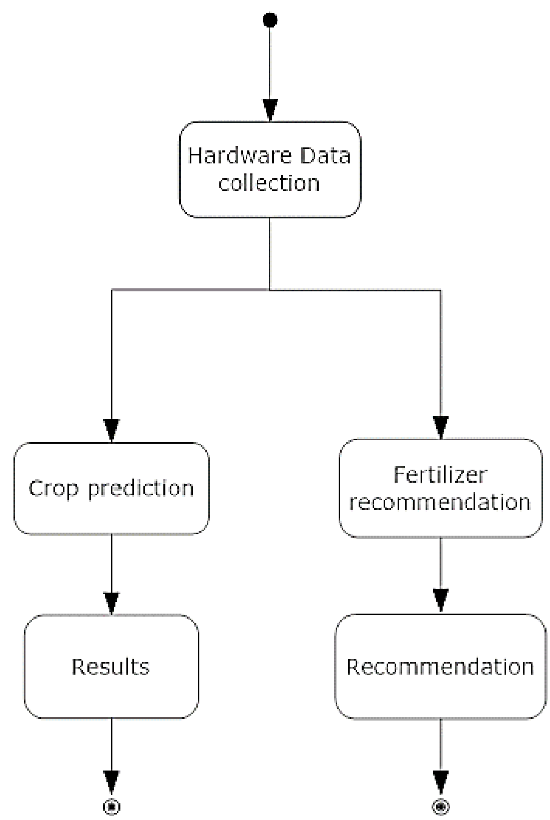


Figure 4.6: State Chart Diagram

4.8 Activity Diagram

The general work-flow of the planner can be graphically represented in an activity diagram. Figure 4. shows how user will use the system and the step-by-step process they will go through as they progress through the site. The diagram shows the workflow for all average user. The user is then able to interact with selected modules, or open new modules. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another

Purpose of Activity Diagrams

The basic purpose of activity diagrams is similar to other UML diagrams. It captures the dynamic behavior of the system. Other UML diagrams are used to show the message flow from one object to another but the activity diagram is used to show message flow from one activity to another.

Activity is a particular operation of the system. Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part.

It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single.

The purpose of an activity diagram can be described as:

- Draw the activity flow of a system.
- Describe the sequence from one activity to another.
- Describe the parallel, branched and concurrent flow of the system

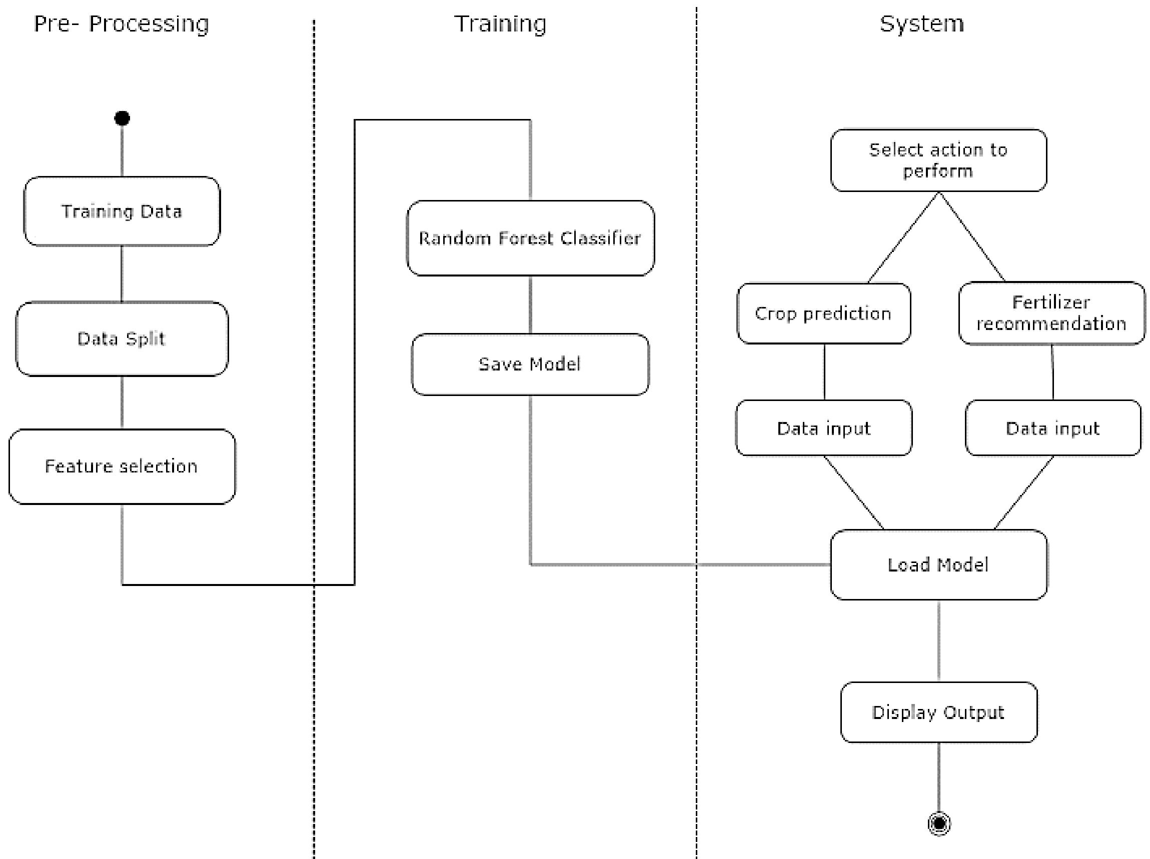


Figure 4.7 : Activity Diagram

5. IMPLEMENTATION

5.1 Implementation Strategy

Crop prediction and fertilizer recommendation systems have become increasingly popular in recent years due to their potential to help farmers increase crop yields and optimize resource utilization. These systems use various data sources such as weather, soil, and crop growth data to make accurate predictions about crop yields and recommend appropriate fertilizers. The advancements in machine learning algorithms and data analysis techniques have made the implementation of such systems more accessible. Open-source machine learning libraries like TensorFlow, Keras, and PyTorch have made it easier to build accurate crop prediction models. These models use various techniques such as neural networks, decision trees, and regression to analyze crop data and make predictions. By analyzing crop data from previous years, these models can make predictions about future crop yields with high accuracy. Additionally, these models can recommend appropriate fertilizers based on the soil and crop data. Mobile applications and web interfaces can be developed to make these systems accessible to farmers. With these interfaces, farmers can input their crop data and receive recommendations for fertilizers and predicted crop yields. These interfaces can also be integrated with weather data to provide real-time updates and recommendations.

The use of crop prediction and fertilizer recommendation systems can provide significant benefits to farmers. By optimizing resource utilization, farmers can save money on fertilizers and increase crop yields. Additionally, accurate crop predictions can help farmers make informed decisions about crop planning and management. However, the implementation of these systems requires careful consideration of the data sources and machine learning algorithms used. The accuracy of the predictions and recommendations relies heavily on the quality of the data used. Therefore, it is crucial to ensure that the data sources used are reliable and accurate. Additionally, the algorithms used must be carefully chosen and tested to ensure accurate predictions and recommendations.

5.2 Hardware Platform Used

In this project, a computer with sufficient processing power is needed. This project also requires IoT sensors and hardware setup.

- Laptop/PC
- System: Intel Processor i3/i5/i7 or AMD processors
- RAM
- Hard Disk: 1GB
- Arduino Uno
- Soil NPK sensor
- Soil Moisture Sensor

5.3 Software Platform Used

- Flask
- Jinja2
- Numpy
- Pandas
- Scikit-learn
- Scipy
- SQLAlchemy
- Pycharm IDE

5.4 Hardware Specification

In this project the system needed to collect real time data from soil sample which needs be captured over period of time and submitted to a system for predictions. Hence we used following IoT components:

5.4.1 Arduino Uno

Arduino Uno is a popular microcontroller board based on the ATmega328P microcontroller. It is an open-source electronics platform that enables you to create interactive electronic projects with ease. It was developed by Arduino LLC and was first released in 2010.

The ATmega328P microcontroller is the brain of the Arduino Uno board, which runs at a clock speed of 16 MHz. The board has 14 digital input/output (I/O) pins, out of which 6 can be used as PWM (Pulse Width Modulation) outputs, 6 analog input pins, and a 16 MHz quartz crystal. It also has a USB port for serial communication and programming.

The board is designed in a way that makes it easy to use, even for beginners. It has a series of headers that can be connected to other components, such as sensors, actuators, and other electronic modules. The board can be powered by either a USB cable or an external power supply, which can be up to 12V.

The Arduino Uno board can be programmed using the Arduino Integrated Development Environment (IDE), which is a software tool used to write and upload code to the board. The IDE is available for Windows, Mac OS X, and Linux platforms, and it supports C/C++ programming languages. The code written in the IDE is compiled and uploaded to the board using a USB cable.

One of the main advantages of the Arduino Uno board is its flexibility. It can be used to build a wide range of projects, such as robots, home automation systems, environmental monitoring systems, and many others. The board is also compatible with a wide range of shields, which are plug-in modules that add specific functionality to the board, such as WiFi connectivity, GPS, or motor control.

Overall, the Arduino Uno board is a powerful and versatile microcontroller board that provides a simple and easy-to-use platform for building electronic projects. Its popularity and large community of users make it an excellent choice for both beginners and advanced users who want to experiment with electronics and programming.

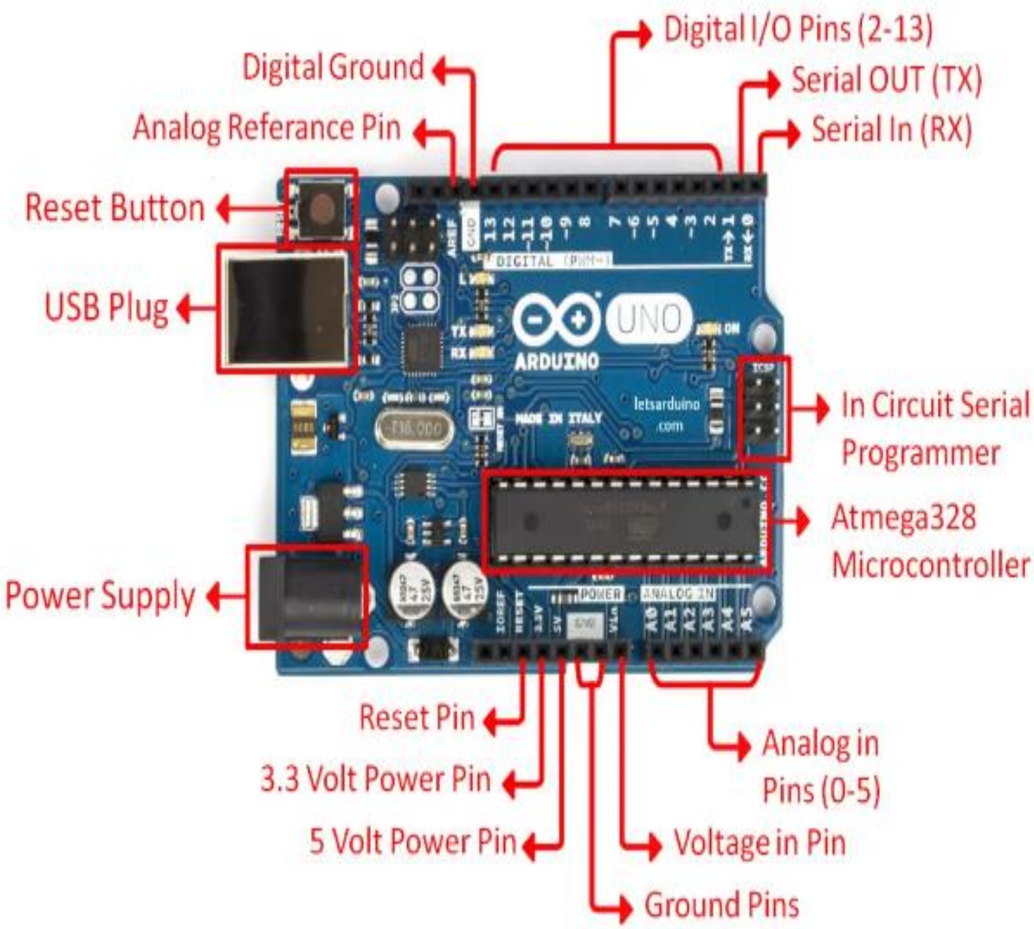


Figure 5.1: Pin diagram of Arduino Uno

The Arduino Uno board has a number of key features, including:

- 14 Digital Input/Output Pins: These pins can be used to interface with a variety of digital devices, such as sensors, LEDs, and switches.
- 6 Analog Input Pins: These pins can be used to measure analog signals, such as light or temperature.
- 16 MHz Clock Speed: This clock speed allows the microcontroller to run code quickly and efficiently.
- USB Connection: This connection allows the board to be programmed and powered

through a computer.

- Power Jack: This jack allows the board to be powered by an external power supply.
- Reset Button: This button resets the board.

5.4.2 Soil NPK Sensor

Soil NPK sensors are devices that measure the levels of three key nutrients in soil: nitrogen (N), phosphorus (P), and potassium (K). These nutrients are essential for plant growth and are often added to soil as fertilizers. Soil NPK sensors are useful for determining the nutrient status of the soil and can help farmers and gardeners make informed decisions about the type and amount of fertilizer to apply.

Soil NPK sensors typically use spectroscopy, electrochemical, or ion-selective electrode technologies to measure nutrient levels. Spectroscopy sensors use light to analyze the chemical composition of the soil, while electrochemical sensors use electrical signals to detect nutrient levels. Ion-selective electrode sensors use a special membrane to selectively detect the ions of interest.

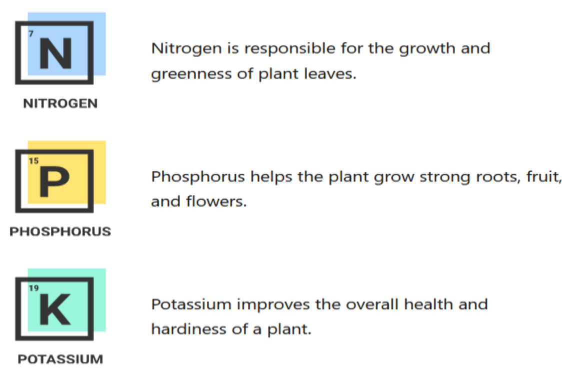


Figure 5.2: Importance of Nutrients

The sensors are usually inserted into the soil at a specific depth and left in place for a certain period of time to allow for accurate measurement of nutrient levels. The sensors may be connected to a data logger or readout device that displays the nutrient levels in real-time. Some sensors also have the ability to transmit data wirelessly, allowing for remote monitoring and control.

Accurate measurement of soil nutrient levels is important for optimal plant growth and crop yield. If soil nutrient levels are too low, plants may not be able to grow properly or produce an abundant yield. Conversely, if soil nutrient levels are too high, it can lead to over-fertilization, which can be harmful to the environment and waste valuable resources. By using soil NPK sensors, farmers and gardeners can determine the nutrient status of the soil and adjust their fertilizer application accordingly. This can help improve crop yield and quality, reduce fertilizer use and costs, and minimize the environmental impact of fertilizer runoff.

Figure 5.3: Soil NPK sensor

Working of NPK Sensor

The sensor includes a stainless steel probe that is rust-proof, electrolytic resistant, and salt-alkali resistant. It can therefore be used with any type of soil, including alkaline soil, acid soil, substrate soil, seedling bed soil, and coconut bran soil. Soil NPK sensors work by measuring the levels of nitrogen (N), phosphorus (P), and potassium (K) in the soil. These nutrients are essential for plant growth and are often added to soil as fertilizers. The sensor operates on 5-30V and consumes very little power. According to the datasheet, it is capable of measuring nitrogen, phosphorus, and potassium with a resolution of up to 1 mg/kg (mg/l).

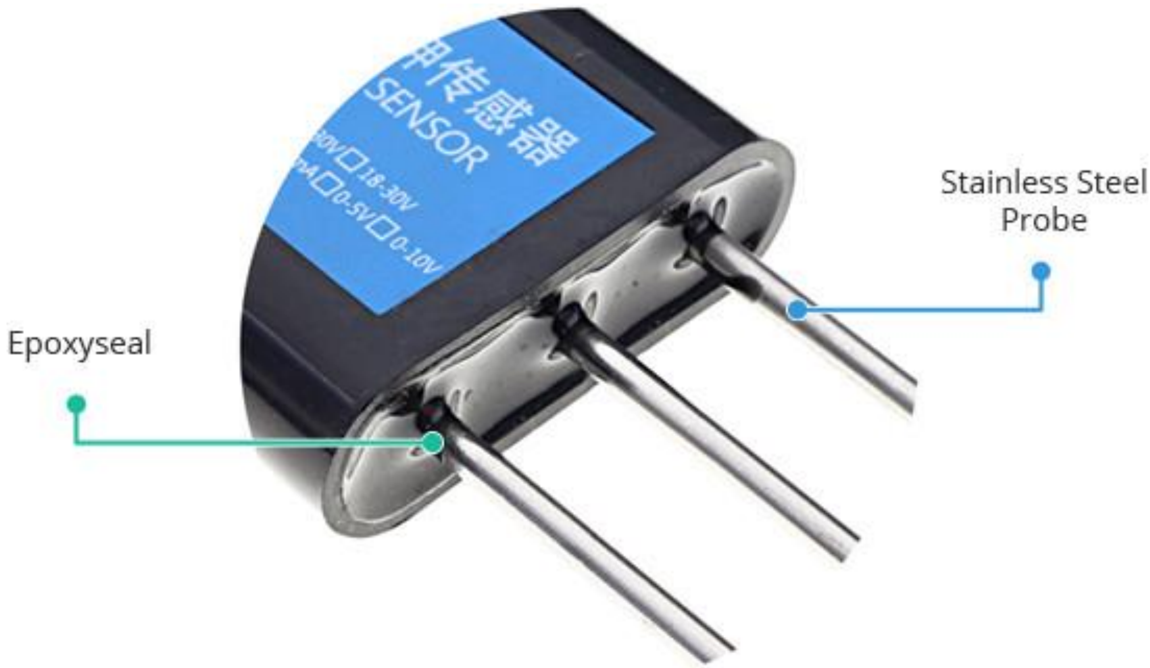


Figure 5.4: Probes of NPK sensor

The probe is sealed to the body with high-density epoxy resin to prevent moisture from entering the body.

The best part is that the sensor has an IP68 rating, which means it is protected against dust and moisture, allowing it to operate normally for a very long time. To be used effectively over long distances, the sensor features the RS485 communication interface and supports the standard Modbus-RTU communication protocol. It should be noted that the sensor cannot be used with an Arduino directly. To communicate with Arduino, you'll need an RS-485 transceiver module that converts a UART serial stream to RS-485.

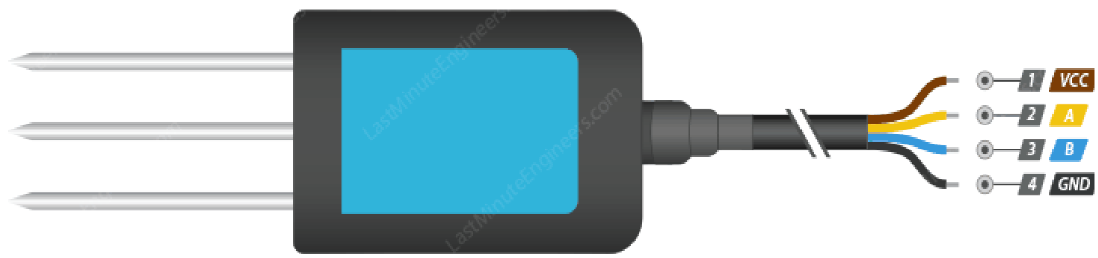


Figure 5.5: NPK Sensor Pin diagram

VCC is the VCC pin. Connects to 5V – 30V.

A is a differential signal that is connected to the A pin of the MAX485 Modbus Module.

B is another differential signal that is connected to the B pin of the MAX485 Modbus Module.

[redacted] is the Ground pin.

There are different types of soil NPK sensors, but most use either spectroscopy, electrochemical, or ion-selective electrode technologies to measure nutrient levels.

Spectroscopy sensors work by shining light on the soil and analyzing the reflected light to determine the chemical composition of the soil. The reflected light contains information about the wavelengths of light that are absorbed and scattered by different soil components, including nutrients such as N, P, and K. By analyzing the reflected light, the sensor can estimate the nutrient levels in the soil. Electrochemical sensors work by using electrodes that are placed in the soil to detect electrical signals generated by the nutrients. The electrodes are connected to a circuit that measures the electrical signals and converts them into nutrient levels.

Ion-selective electrode sensors work by using a special membrane that selectively detects the ions of interest. The membrane is placed in contact with the soil and the ions in the soil diffuse through the membrane to the sensor. The sensor then measures the concentration of the ions in the soil and converts it into nutrient levels.

Wiring a Soil NPK Sensor to an Arduino

The NPK sensor cannot be used directly with an Arduino. To communicate with Arduino, you’ll need an RS-485 transceiver module that converts a UART serial stream to RS-485, such as the one shown below.

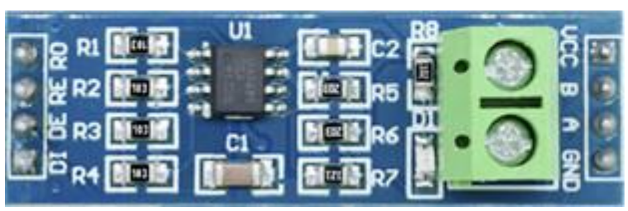


Figure 5.6: RS-485 transceiver module

let’s get to the wiring.

The soil NPK Sensor has four wires. The power wire is brown and should be connected to the 5V-30V power supply. The ground wire is black and should be connected to a common ground.

The yellow wire of the NPK sensor should be connected to the RS485 module’s A pin, and the blue wire should be connected to the RS485 module’s B pin.

Connect the RS485 module's R0 and DI pins to the Arduino's digital pins 2 and 3, respectively. These digital pins will be used as virtual RX and TX serial lines.

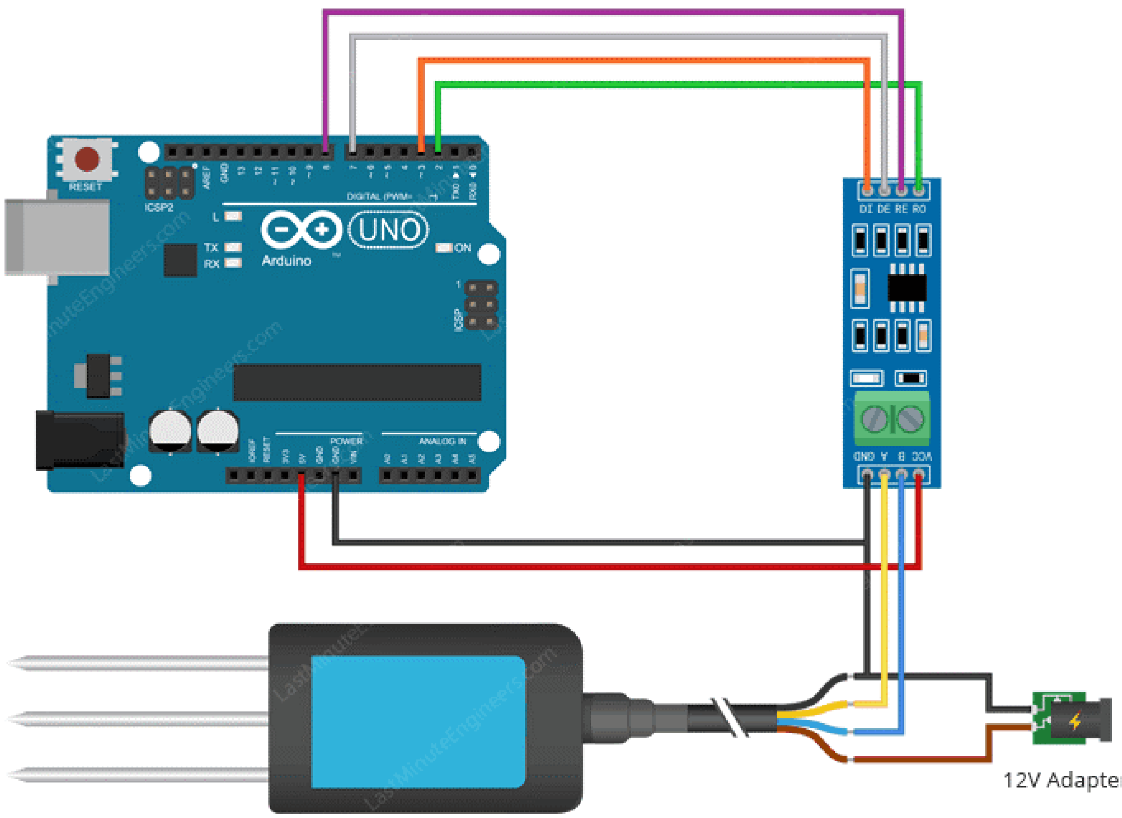


Figure 5.7: Circuit diagram for NPK sensor

5.4.3 Soil Moisture Sensor

Soil moisture sensors are devices that measure the amount of moisture present in the soil. They are commonly used in agriculture, horticulture, and environmental monitoring to help manage irrigation, prevent overwatering or underwatering, and optimize plant growth. There are several types of soil moisture sensors, but the most common ones are capacitance-based sensors and resistive sensors.

Capacitance-based sensors measure the changes in the electrical capacitance between two or more electrodes as the soil moisture content changes. Resistive sensors, on the other hand, measure the resistance of a material (usually a ceramic or gypsum block) that is in contact with the soil.

The basic principle behind the operation of a soil moisture sensor is that the electrical properties of soil change as it becomes wet or dry. When soil is wet, it has a higher electrical conductivity due to the presence of dissolved salts and minerals. Conversely, dry soil has a lower conductivity due to the absence of water.

Soil moisture sensors typically consist of a probe that is inserted into the soil at a specific depth. The probe is connected to a data logger or readout device that displays the moisture content in real-time. Some sensors also have the ability to transmit data wirelessly, allowing for remote monitoring and control.

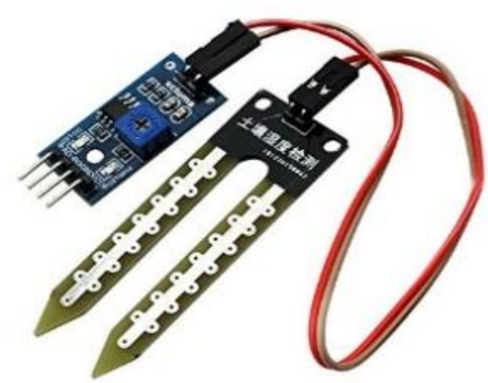


Figure 5.8: Soil Moisture Sensor

Accurate soil moisture measurement is critical for proper irrigation management, as both overwatering and underwatering can negatively impact plant growth and crop yields. By using soil moisture sensors, farmers and gardeners can make informed decisions about when to water their crops or plants, ensuring that they receive just the right amount of water for optimal growth and health. Additionally, soil moisture sensors can help conserve water resources by preventing excessive irrigation and runoff, which can lead to soil erosion and water pollution.

Working of Soil Moisture Sensor

The soil moisture sensor operates in a straightforward manner.

The fork-shaped probe with two exposed conductors acts as a variable resistor (similar to a potentiometer) whose resistance varies with the soil's moisture content.

This resistance varies inversely with soil moisture:

- The more water in the soil, the better the conductivity and the lower the resistance.
- The less water in the soil, the lower the conductivity and thus the higher the resistance.

The sensor produces an output voltage according to the resistance, which by measuring we can determine the soil moisture level.

A typical soil moisture sensor consists of two parts.

1. The Probe

The sensor includes a fork-shaped probe with two exposed conductors that is inserted into the soil or wherever the moisture content is to be measured.

As previously stated, it acts as a variable resistor, with resistance varying according to soil moisture.

Figure 5.9: Probes

2. The Module

In addition, the sensor includes an electronic module that connects the probe to the Arduino.

The module generates an output voltage based on the resistance of the probe, which is available at an Analog Output (AO) pin.

The same signal is fed to an LM393 High Precision Comparator, which digitizes it and makes it available at a Digital Output (DO) pin.

Figure 5.10: Moisture sensor module

The module includes a potentiometer for adjusting the sensitivity of the digital output (DO). We can use it to set a threshold, so that when the soil moisture level exceeds the threshold, the module outputs LOW otherwise HIGH.

This setup is very useful for triggering an action when a certain threshold is reached. For example, if the moisture level in the soil exceeds a certain threshold, you can activate a relay to start watering the plant.

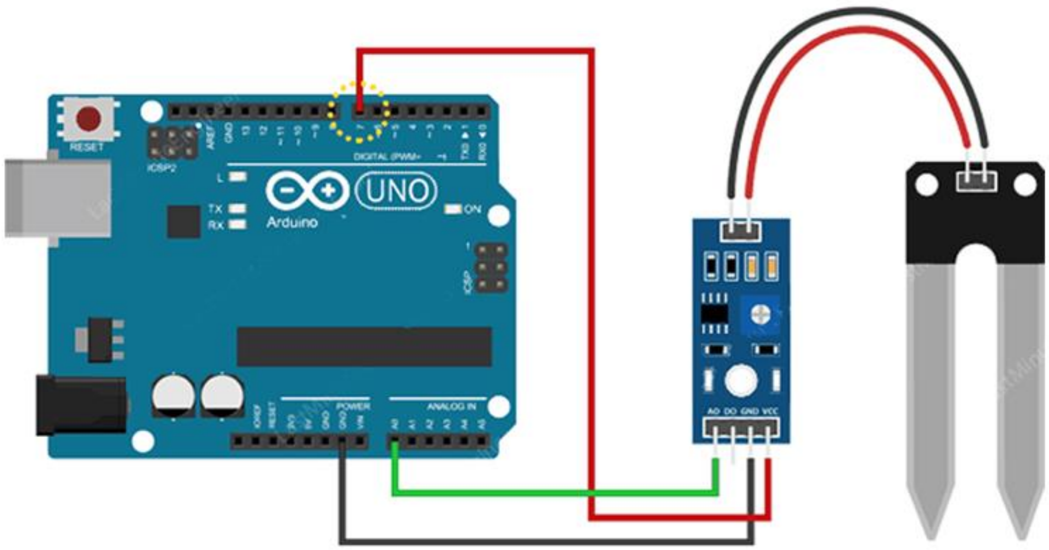


Figure 5.11: Soil Moisture Sensor Circuit

5.5 Software Specification

In this project we used various python libraries to create machine learning model, also used scripting languages like HTML, CSS for web application development and web framework which is flask uses python.

5.5.1 Arduino IDE

The Arduino Integrated Development Environment (IDE) is a software platform used to write, compile, and upload code to Arduino microcontrollers. It is a cross-platform application that works on Windows, Mac, and Linux operating systems.

The Arduino IDE provides a user-friendly interface that allows users to write code in a simplified version of the C++ programming language. The IDE includes a text editor with syntax highlighting, code completion, and error highlighting, making it easy for users to write and debug their code.

In addition to the code editor, the Arduino IDE includes a compiler that converts the written code into machine-readable instructions, and an uploader that transfers the compiled code to the Arduino microcontroller.

The Arduino IDE also includes a range of built-in functions and libraries that make it easier to interact with the hardware components connected to the Arduino board, such as sensors, actuators, and displays.

Overall, the Arduino IDE provides an easy-to-use platform for beginners and advanced users to develop and upload code to Arduino microcontrollers, making it an ideal tool for prototyping and developing various projects.

5.5.2 Python libraries used

1. **Flask:** Flask is a micro web framework for building web applications in Python. It was created by Armin Ronacher and is classified as a micro framework because it does not require particular tools or libraries. Flask is easy to learn and allows developers to quickly prototype and build web applications with minimal boilerplate code.

Flask provides the basics for building a web application, such as routing, request and response handling, and template rendering. However, it does not come with features such as database abstraction layers, form validation, or authentication, which can be added through extensions or third-party libraries. One of the strengths of Flask is its flexibility. It does not impose any particular structure or architecture on the application, allowing developers to choose the components they want to use and how they want to organize their code. Flask also has a vibrant ecosystem of extensions and plugins, which can be used to add functionality to the application.

Overall, Flask is a lightweight and versatile web framework that is suitable for building small to medium-sized web applications or prototypes. Its simplicity and flexibility make it a popular choice among Python developers.

2. **Jinja2:** Jinja2 is a popular templating language for Python web applications. It was inspired by Django's template engine and was created by Armin Ronacher, the same developer who created the Flask web framework.

   Jinja2 allows developers to separate the presentation layer from the business logic layer of a web application. This means that developers can focus on building the backend functionality of an application without worrying about how it will be presented to the user.

   Jinja2 templates are simple text files that can contain placeholders, called variables, which are replaced with dynamic content at runtime. Templates can also include control structures, such as if/else statements and loops, which allow developers to dynamically generate the content of a web page based on the application's data. Jinja2 is highly extensible and can be customized to fit the needs of a specific application. Developers can define their own filters, functions, and global variables to be used in templates, allowing them to encapsulate complex logic and reuse it across multiple templates.

   Jinja2 is widely used in Python web frameworks, including Flask, Django, and Pyramid. It is known for its speed, flexibility, and ease of use, and has a large community of developers contributing to its development and maintenance. Overall, Jinja2 is a powerful and versatile templating language that is essential for building dynamic and responsive web applications in Python.

3. **Numpy:** NumPy is a Python library that provides support for large, multi-dimensional arrays and matrices, as well as a variety of mathematical operations to be performed on these arrays. It is one of the core libraries used in scientific computing with Python and is a fundamental tool for numerical analysis, scientific simulations, and data science. NumPy is built on top of the C programming language, which makes it fast and efficient when working with large datasets. It provides an extensive set of mathematical functions, including linear algebra, Fourier transform, random number generation, and statistical analysis, which can be applied to arrays of any dimensionality.

   One of the key features of NumPy is its ability to perform vectorized operations, which allow mathematical operations to be performed on entire arrays at once, instead of looping through each element of the array. This greatly improves the performance of mathematical computations and makes it possible to handle large datasets with ease.In addition to its support for arrays and matrices, NumPy also provides tools for working with other data types, such as structured arrays, masked arrays, and broadcasting arrays. It also has functions for integrating with other libraries, such as SciPy, Pandas, and Matplotlib, which makes it a powerful tool for scientific computing and data analysis.

   Overall, NumPy is a powerful and essential library for anyone working with scientific computing or data analysis in Python. Its efficient handling of large datasets, extensive mathematical functions, and integration with other libraries make it a versatile and indispensable tool for a wide range of applications.

4. **Pandas:** Pandas is a popular open-source library for data manipulation and analysis in Python. It provides high-performance, easy-to-use data structures and data analysis tools

for working with structured data, such as tables or spreadsheets. The two primary data structures in Pandas are the Series and DataFrame objects.

The Series is a one-dimensional array-like object that can hold any data type, such as integers, floats, or strings. The DataFrame is a two-dimensional table-like object that is used to store and manipulate data in rows and columns. Both the Series and DataFrame objects are highly flexible and can be indexed and sliced in a variety of ways.

Pandas provides a wide range of data manipulation and analysis functions, such as data cleaning, merging, grouping, and aggregation. It also supports advanced data analysis tasks, such as time-series analysis, data visualization, and machine learning. One of the key features of Pandas is its ability to handle missing data. It provides functions for filling, dropping, and interpolating missing data, making it easy to clean and prepare data for analysis.

Pandas is widely used in data science, finance, and other industries for data cleaning, exploration, and analysis. It has a large and active community of users and contributors, and is constantly being updated and improved with new features and functionality.

Overall, Pandas is a powerful and flexible tool for working with structured data in Python. Its rich set of functions and easy-to-use data structures make it a must-have library for anyone working with data analysis and manipulation in Python.

5. **Scikit-learn:** Scikit-learn is a popular open-source machine learning library for Python. It provides a wide range of tools and algorithms for data analysis, data preprocessing, and machine learning, making it one of the most widely used libraries in the field of data science.

   Scikit-learn is built on top of other Python libraries, such as NumPy, SciPy, and Pandas, and provides a unified interface for working with a variety of machine learning algorithms, including classification, regression, clustering, and dimensionality reduction.

One of the key features of Scikit-learn is its ability to work with both small and large datasets. It provides efficient algorithms for handling large datasets, such as stochastic gradient descent and mini-batch gradient descent, which can scale to handle datasets with millions or even billions of examples.

Scikit-learn also provides a wide range of data preprocessing and feature engineering functions, including data normalization, feature scaling, and feature selection. These functions allow users to transform and prepare their data for machine learning algorithms. In addition to its core machine learning algorithms, Scikit-learn also provides a variety of tools for model selection and evaluation, including cross-validation, grid search, and metrics for evaluating model performance.

Scikit-learn is widely used in data science, machine learning, and artificial intelligence, and has a large and active community of users and contributors. It is constantly being updated and improved with new algorithms and features, making it a must-have library for anyone working with machine learning in Python.

Overall, Scikit-learn is a powerful and versatile library for machine learning in Python. Its ease of use, scalability, and extensive set of tools and algorithms make it an essential tool for anyone working with data science or machine learning applications.

6. **Scipy:** SciPy is a Python library for scientific and technical computing that provides a wide range of functions for numerical optimization, integration, interpolation, signal processing, linear algebra, and more. It is built on top of NumPy and is one of the core libraries used in scientific computing with Python.

   One of the key features of SciPy is its extensive set of mathematical functions, which can be used for a wide range of scientific and technical applications. It provides functions for solving differential equations, performing Fourier transforms, calculating statistical distributions, and more.

   SciPy also provides tools for optimization, including nonlinear optimization, least-squares minimization, and root finding. These tools can be used to find the optimal solution for a wide range of problems in science, engineering, and finance.In addition to its mathematical and optimization functions, SciPy also provides tools for signal processing, image processing, and machine learning. It has functions for filtering, smoothing, and transforming signals and images, as well as algorithms for clustering, classification, and regression.

SciPy is widely used in scientific computing, data science, and machine learning, and has a large and active community of users and contributors. It is constantly being updated and improved with new features and functionality, making it a must-have library for anyone working with scientific and technical computing in Python.

Overall, SciPy is a powerful and essential library for anyone working in scientific computing, data science, or machine learning with Python. Its extensive set of mathematical functions, optimization tools, and signal processing functions make it a versatile and indispensable tool for a wide range of applications.

7. **SQLAlchemy:** SQLAlchemy is a popular open-source SQL toolkit and Object-Relational Mapping (ORM) library for Python. It provides a high-level, Pythonic interface for working with relational databases, allowing developers to work with databases in a more natural and intuitive way.

   SQLAlchemy provides a rich set of tools for interacting with databases, including tools for creating and managing database schemas, executing queries, and managing transactions. It also provides support for a wide range of database engines, including PostgreSQL, MySQL, SQLite, and Microsoft SQL Server, among others.

   One of the key features of SQLAlchemy is its Object-Relational Mapping (ORM) system. The ORM system allows developers to map database tables to Python classes, and map database columns to Python class attributes. This allows developers to work with databases in a more natural and intuitive way, and to write database queries using Python code rather than SQL.

5.6 Deployment Diagram

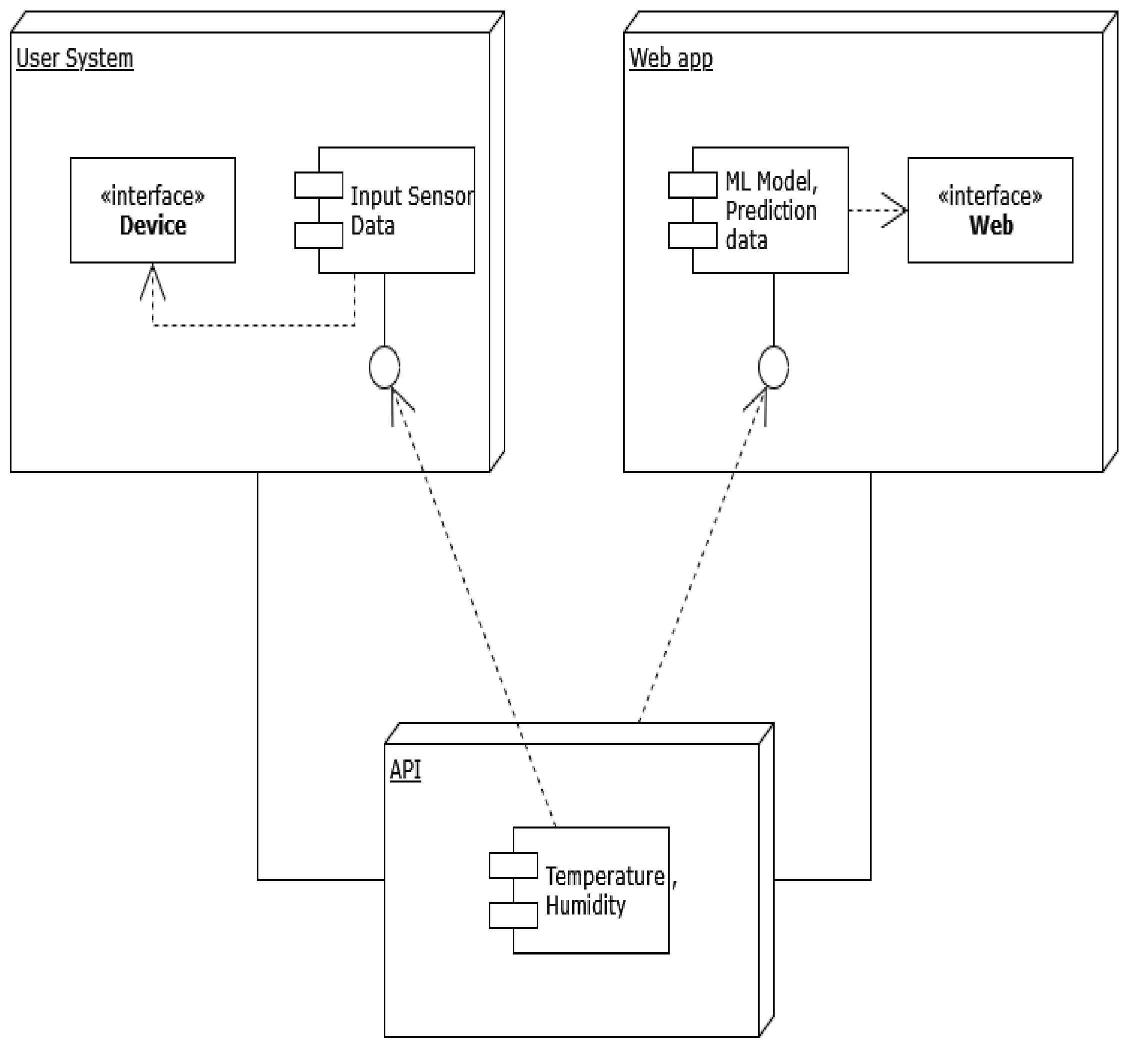


Figure 5.12 Deployment Diagram

Deployment diagrams are used to visualize the topology of the physical components of a system, where the software components are deployed as shown in below figure 5.1. Deployment diagrams are used to describe the static deployment view of a system. Deployment diagrams consist of nodes and their relationships.

5.7 Implementation Level Details

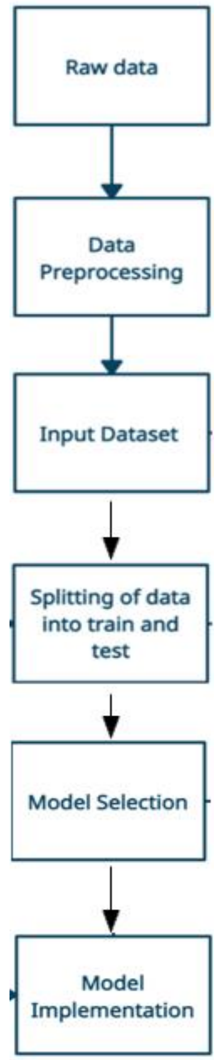


Figure 5.13: Implementation Model

5.7.1 Dataset Details

For making the model efficient for real-time prediction. We have gathered the data from different available data sets like from government databases which are available on data.gov.in, Kaggle. Further, we have searched dataset which is more efficient for the project. As it is easily available as well as because of its features. It has been broken into two to three different datasets.

The dataset we have used contains 2201 data rows and six columns. The column names are N, P, K, temperature, humidity, and label.

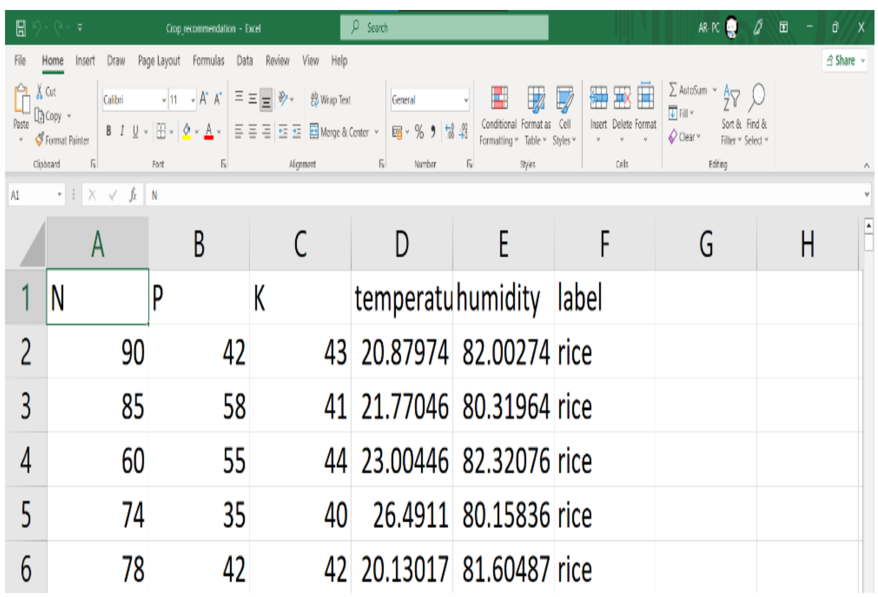
N, P, and K represent the soil nutrients Nitrogen, Phosphorous, and Potassium respectively. These three nutrients are essential for plant growth and are commonly found in fertilizers.

Temperature is the average temperature of the soil in Celsius. Soil temperature is an important factor for crop growth and development as it affects seed germination, nutrient uptake, and disease incidence.

Humidity represents the relative humidity of the air in percentage. Relative humidity is the amount of water vapor present in the air compared to the maximum amount that can be present at a particular temperature.

The label column represents the output or target variable, which could be the crop yield or the crop type.

This dataset can be used for training machine learning models to predict the crop type based on the input features of soil nutrients, temperature, and humidity.



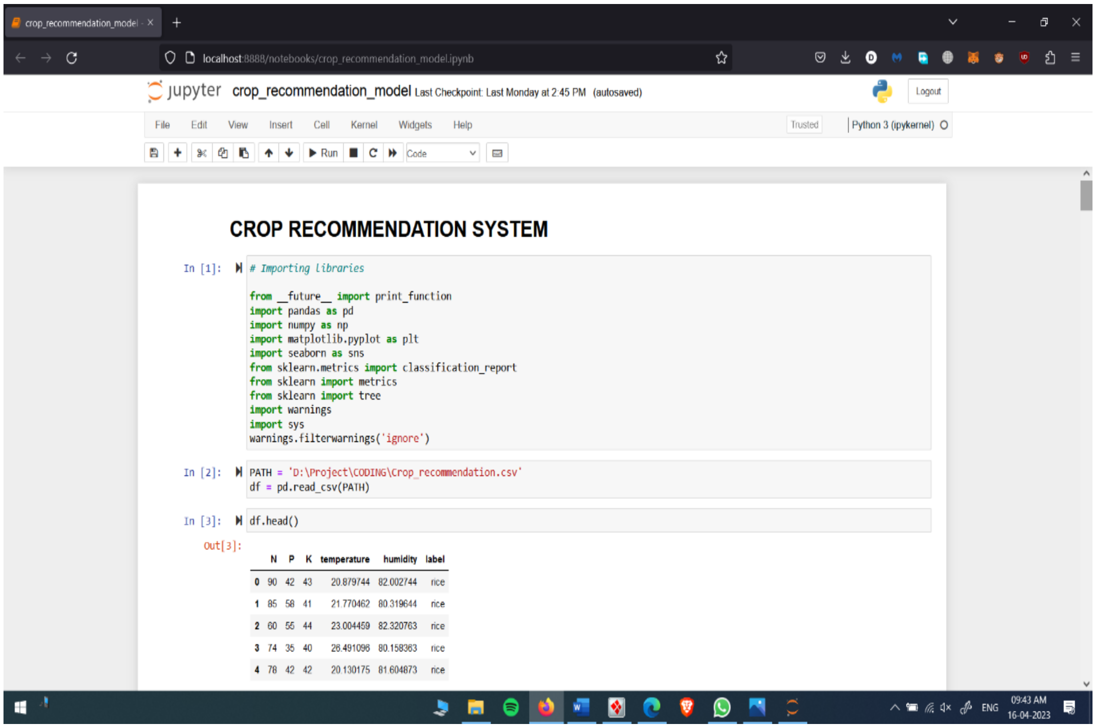
	A	B	C	D	E	F	G	H
1	N	P	K	temperatu	humidity	label		
2	90	42	43	20.87974	82.00274	rice		
3	85	58	41	21.77046	80.31964	rice		
4	60	55	44	23.00446	82.32076	rice		
5	74	35	40	26.4911	80.15836	rice		
6	78	42	42	20.13017	81.60487	rice		

Figure 5.14 : Dataset

5.7.2 Module 2: Pre-processing

Preprocessing is an important step in preparing the data for machine learning algorithms. Here are some preprocessing steps that we applied to the dataset described:

- Missing values: Check if there are any missing values in the dataset. If there are, decide on an appropriate strategy to handle them, such as imputing them with the mean or median of the column, or dropping the rows with missing values.
- Outliers: Check for outliers in the data, which are data points that lie far from the typical range of values for that feature. Outliers can be removed, or their values can be adjusted to the closest typical value.
- Feature scaling: Some machine learning algorithms work better when the features are scaled to have similar ranges. You can use techniques like normalization or standardization to scale the features.
- Encoding categorical variables: If any of the features are categorical variables, such as the label column, they will need to be encoded into numerical values. This can be done using techniques like one-hot encoding.
- Feature selection: Depending on the machine learning algorithm being used, it may be beneficial to select only the most relevant features for the prediction task. Feature selection techniques like correlation analysis or principal component analysis can be used to identify the most important features.
- Splitting the data: Finally, the data can be split into training and testing sets. The training set is used to train the machine learning model, while the testing set is used to evaluate its performance. A common split is to use 80% of the data for training and 20% for testing.



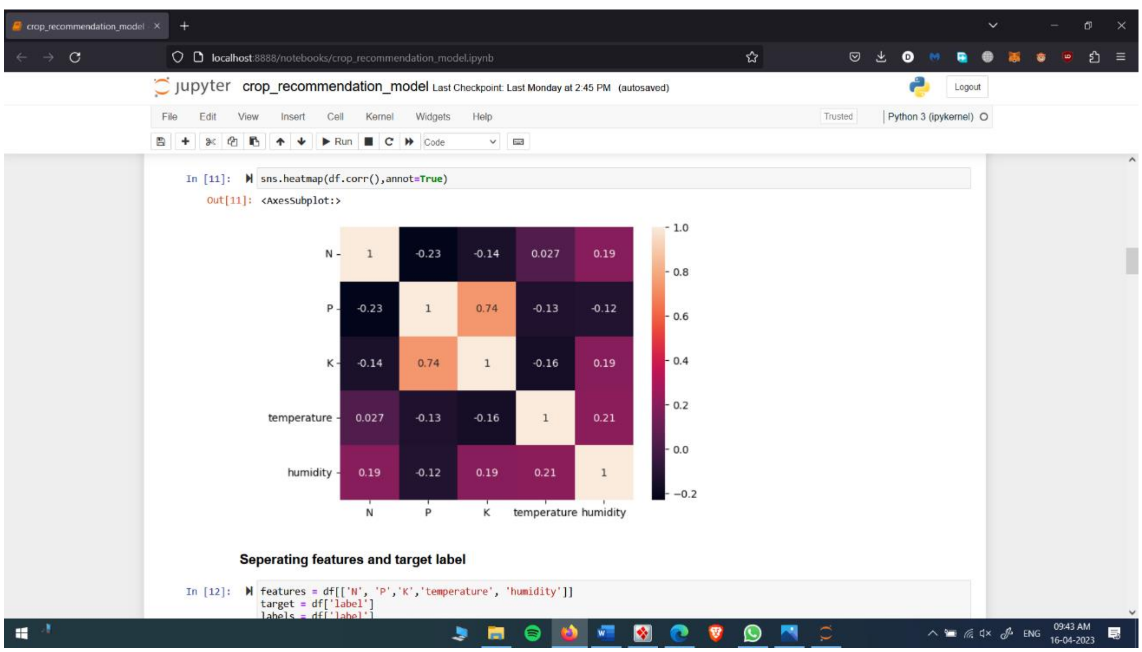


Figure 5.15: Pre-Processing

5.7.3 Module 3: Random Forest Model

Before selecting an algorithm, we need to first assess and compare different options, and then choose the one that best fits our specific dataset. Machine Learning is the best approach for providing a practical solution to the crop selection problem.

There are various ML algorithms that can be used for predicting crop. This study will consider the following algorithms for selection and accuracy comparison:

- Naïve Bayes Classifier: A popular machine learning algorithm for classification tasks is Naive Bayes. It depends on Bayes' hypothesis and accepts that the elements used to characterize the information are autonomous of one another. Given the input features, the probability of each class is calculated by the algorithm, and the class with the highest probability is chosen as the predicted class. It gives an accuracy of 94%.

- Logistic Regression: A classification algorithm called logistic regression is used to predict the probability of a target variable. There are only two classes that can be found in the target binary variable. At the point when strategic relapse is applied to our dataset it has given an accuracy of 87%.

- Decision Tree: The Decision Tree is a method of machine learning that can be used for classification and regression. The model appears as a tree, with every hub addressing a property or component, each branch addressing a choice rule, and each leaf hub meaning a class mark or anticipated esteem. The calculation separates the information into subsets in view of the information highlight values, determined to deliver subsets that are essentially as homogeneous as conceivable with respect to the objective variable. The procedure is recursive and goes on until a stopping point is reached. When applied to the dataset, it has produced an accuracy of 80.0%.

- Random Forest: Based on various data subsets, the algorithm creates decision trees and then makes predictions for each subset. The algorithm improves the system's solution by combining the results from a voting process. The output of Random Forest is more accurate because it trains the data using the bagging technique. RF has provided an accuracy of 95 percent for these data.

We built a Random Forest model which is based on RF(Random Forest Classifier) as we got maximum accuracy and precision with RF. We selected dependent and independent features which will be required for crop prediction and classification and also to recommend fertilizers application using pre-existing values of required nutrients

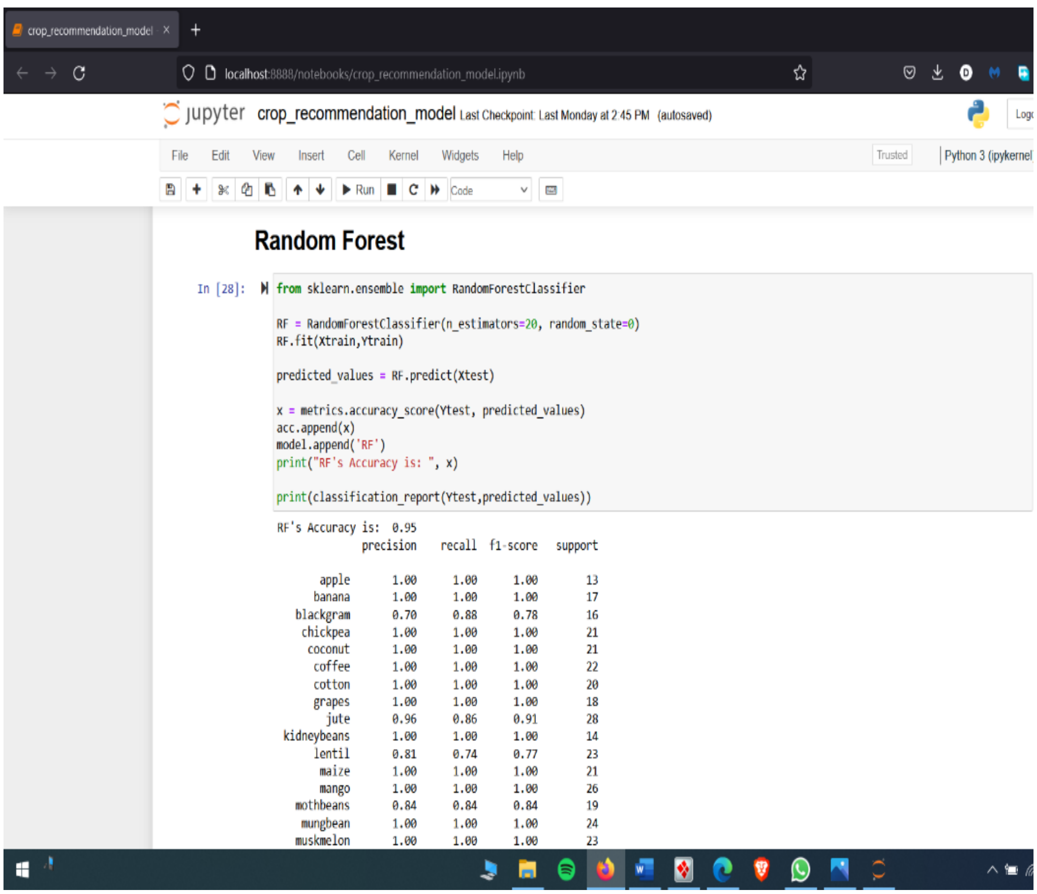


Figure 5.16: Prediction model

Random forest is a popular ensemble learning method used in machine learning and data mining. It is a type of decision tree algorithm that builds multiple decision trees and combines them to make a more accurate prediction. The name "random forest" comes from the fact that each decision tree in the forest is trained using a random subset of the available features and data. The final random forest prediction is made by aggregating the predictions of all the individual trees in the forest. Random forest is widely used for classification and regression tasks and is known for its robustness, accuracy, and ability to handle high-dimensional data.

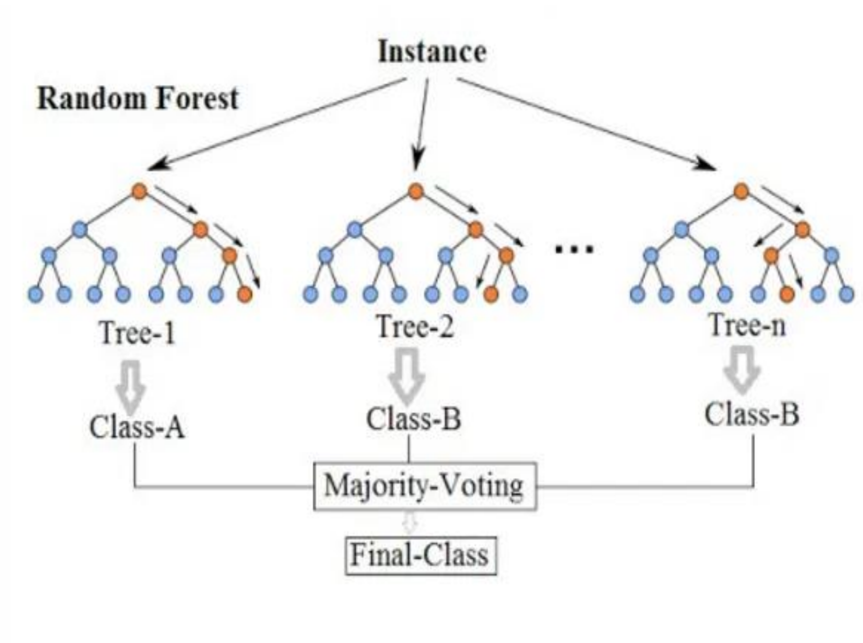


Figure 5.17: Random Forest model simplified

Here is a step-by-step overview of how random forest works:

Data is randomly sampled with replacement from the original dataset to create multiple subsets (or bags) of the data.

A decision tree is built for each subset of data using a random selection of features.

The decision trees are combined to form a random forest. The final prediction is made by taking the mode (for classification) or the mean (for regression) of the predictions of all the individual trees in the forest.

Random forest is designed to reduce the risk of overfitting, which occurs when a model is too complex and fits the training data too closely, but fails to generalize to new data. By building multiple decision trees on different subsets of data and features, random forest is able to capture the patterns and relationships in the data more accurately and reduce the variance in the predictions.

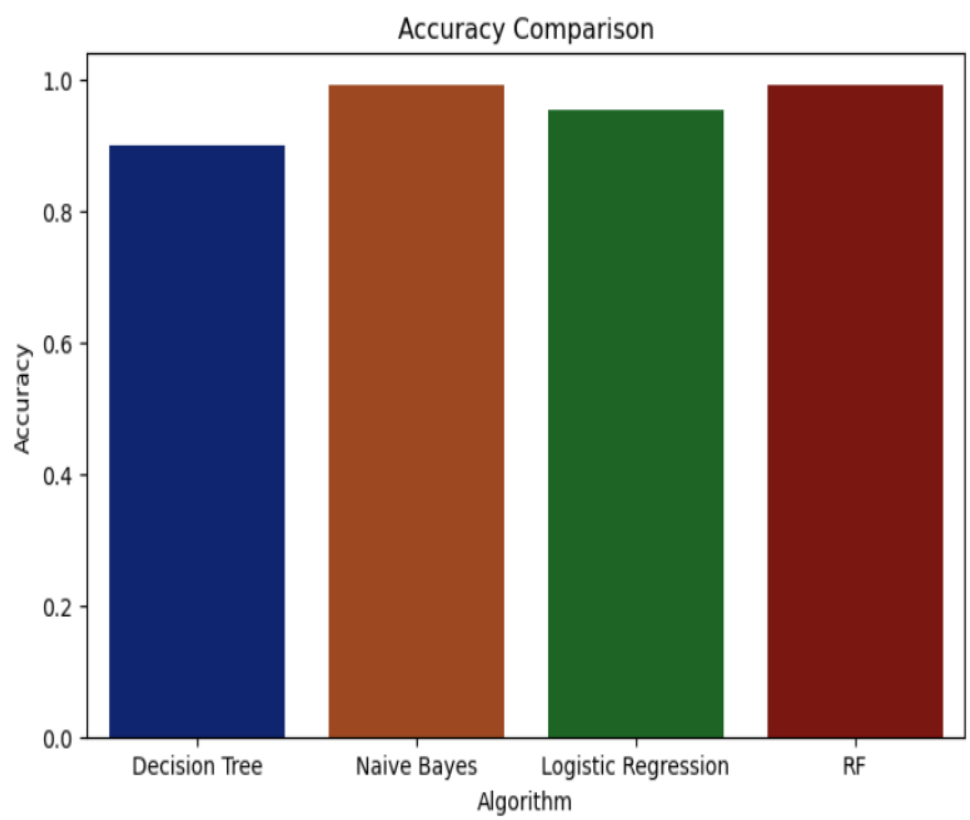


Figure 5.18: Accuracy Comparison of Models

5.8 Testing

GUI testing:

Graphical User Interface (GUI) testing is the one of the mechanism in which user interface developed System Under some graphical rules. GUI testing includes checking various controls- menus, buttons, icons, dialog boxes and windows etc. Proposed system is tested for user inputs against different modules, validations are done.

Unit Testing:

It is the testing of individual software units of the application. It is done after the complexion of an individual unit before integration. Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. This is a structural testing that relies on knowledge of its construction and is invasive. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration Testing: Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfactory, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

System testing: System testing tests a completely integrated system to verify that the system meets its requirements. For example, a system test might involve testing a logon interface, then creating and editing an entry, plus sending or printing results, followed by summary processing or deletion (or archiving) of entries, then logoff.

6. CONCLUSION

This system focuses on the prediction of crop and fertilizers recommendations with the help of machine learning techniques and IoT sensors. Various machine-learning techniques were employed for the computation of precision. The Random Forest classifier was implemented to predict the harvest for a specific soil contains. A mechanism was established to anticipate crop yield based on prior data. The given System assists farmers in making choices regarding which crop to grow on their agricultural land. This system improves the overall efficiency of farmland. By maximizing crop yield rates, this undertaking enhances the Indian economy. From the study in this paper, we concluded that there is still a need of research in the agricultural field to get better accuracy. Using ensemble methods is a good way to ensure better accuracy of the system. Also, if we want to consider only one algorithm for the recommendation system then we can use SVM due to its simple computational requirements. Future work in this field can be geospatial analysis combining and using all the seasonal, soil, weather, temperature, topographical, crop production and economic condition of the farmer into one single model to provide a robust and centralized system for the user to access. Particularly, the study showed that for the previously proposed recommendation systems, factors such as user's investment in farming, number of people for the maintenance and the stretch of land available for cultivation were not considered. These parameters also play a major role in the profits of the farmer which is the basic reason why a farmer would use a crop recommendation system.

FUTURE WORK

Nothing is perfect and complete and there is always a scope of improvement in each and every product. Everything needs to be updated or upgraded on a timely basis to cope up with the current technology. There are many future scopes for this project and hopefully it will emerge into the biggest benefit in the field of artificial intelligence. There were many features we had hoped to integrate into our system, but we were compelled to cut them due to time constraints. However, given more time to work on the product, there are a few changes that we would make.

There is always a scope for enhancements in any developed system, especially when the project build using latest trending technology and has a good scope in future.

- Use of Cloud: Cloud services can be used to provide real-time information on weather patterns and soil moisture levels. This data can be integrated with real-time crop prediction and fertilizer recommendation systems to provide even more accurate predictions and recommendations.

- Integration with blockchain technology: Blockchain technology can be used to create a transparent and secure record of agricultural data. This can help farmers to track the origin of their crops and ensure that they are using sustainable farming practices.

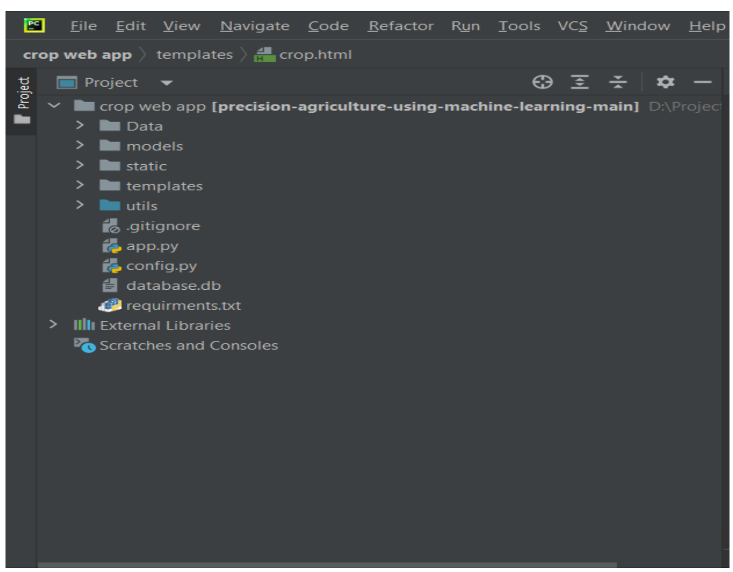
- Use of artificial intelligence: Artificial intelligence techniques such as deep learning can be used to analyze large volumes of data and identify patterns that are not immediately apparent to humans. This can help farmers to optimize their crop management practices and increase yields.

- Integration with mobile apps: Mobile apps can be developed to provide farmers with real-time access to crop predictions and fertilizer recommendations. This will enable farmers to make decisions on the go and respond quickly to changes in weather or soil conditions.

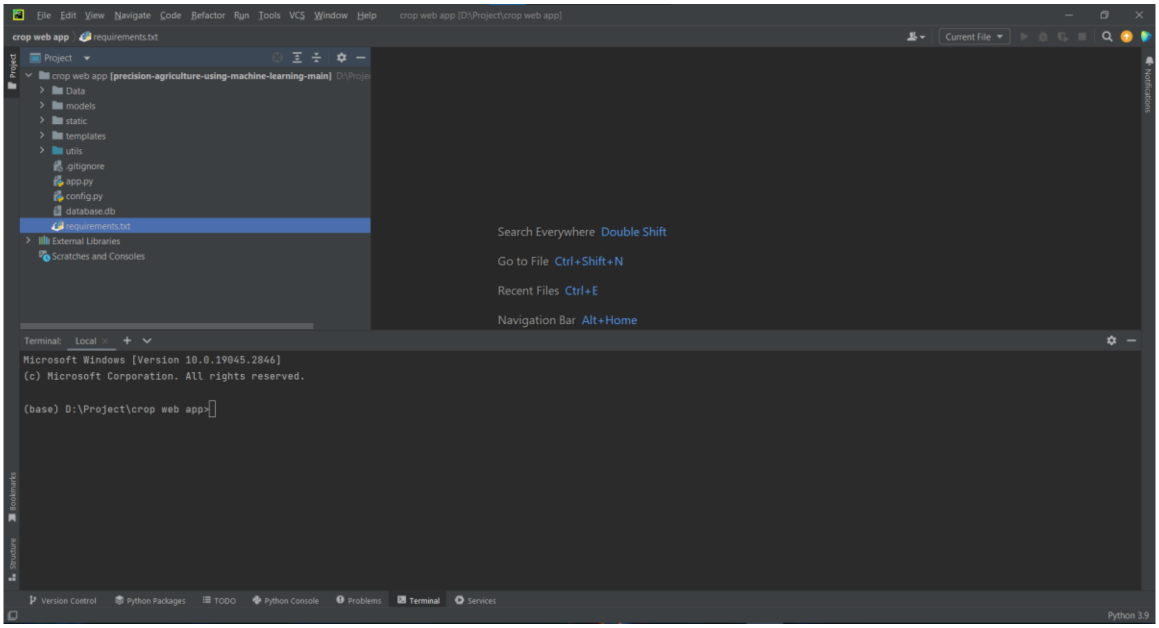
User Manual

1. Open the PyCharm IDE for opening the project.

2: Open the folder in which web application files are stored.

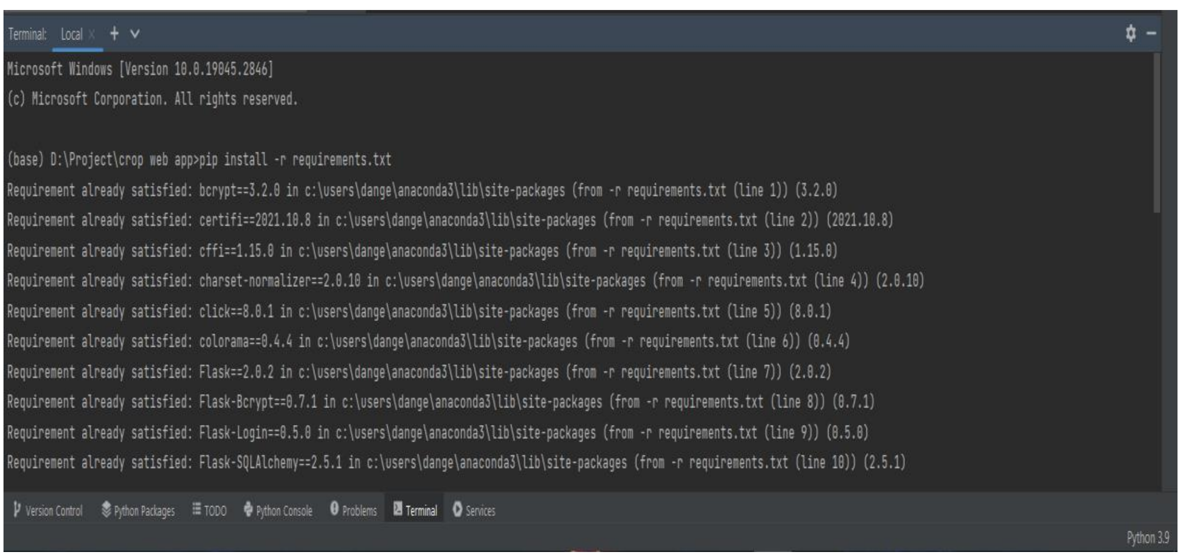


3: After opening folder, Open terminal in PyCharm, and install requirements.



For installing requirements, type in terminal

Pip install -r requirements.txt

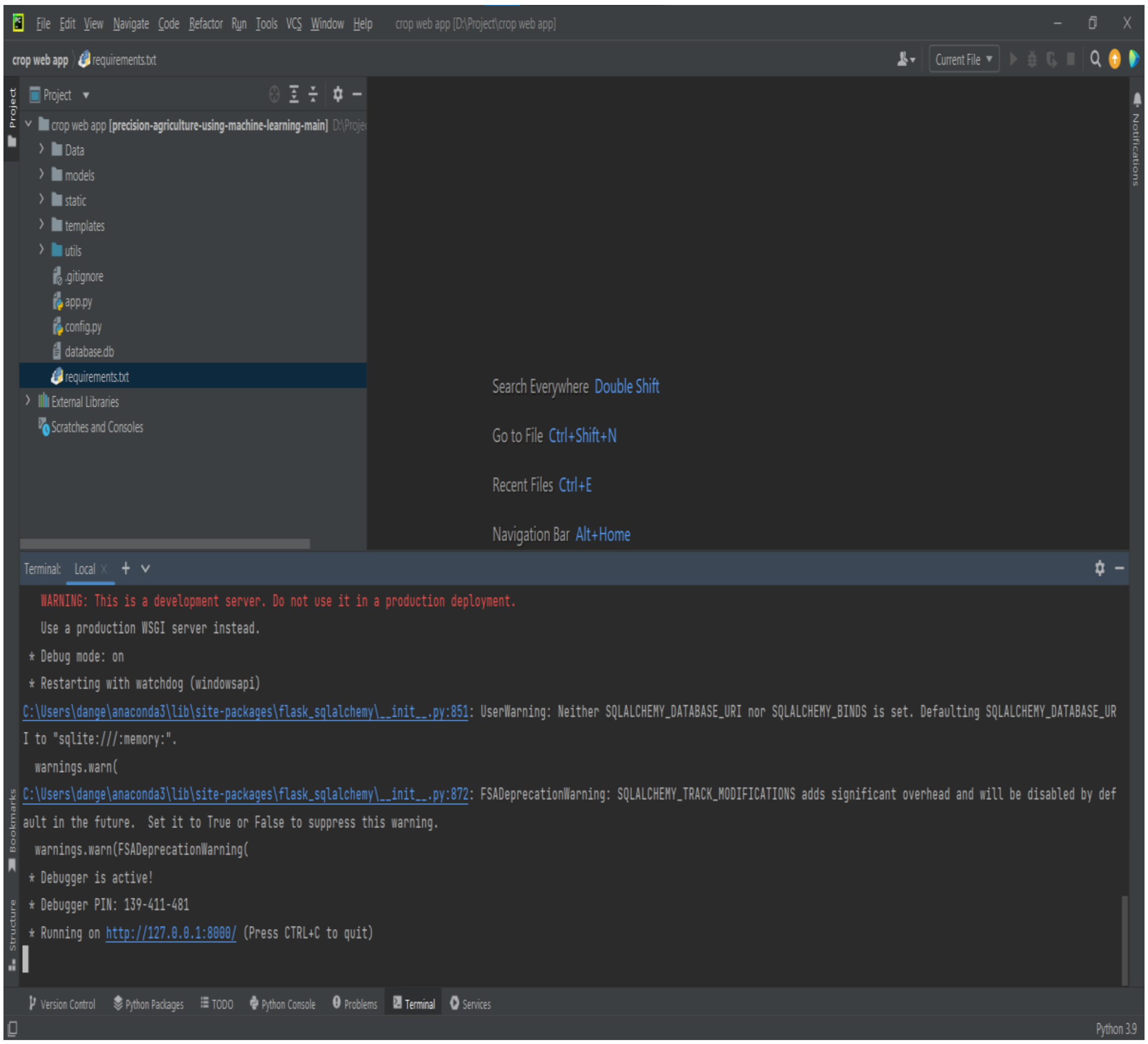


You can see required python libraries getting installed in system.

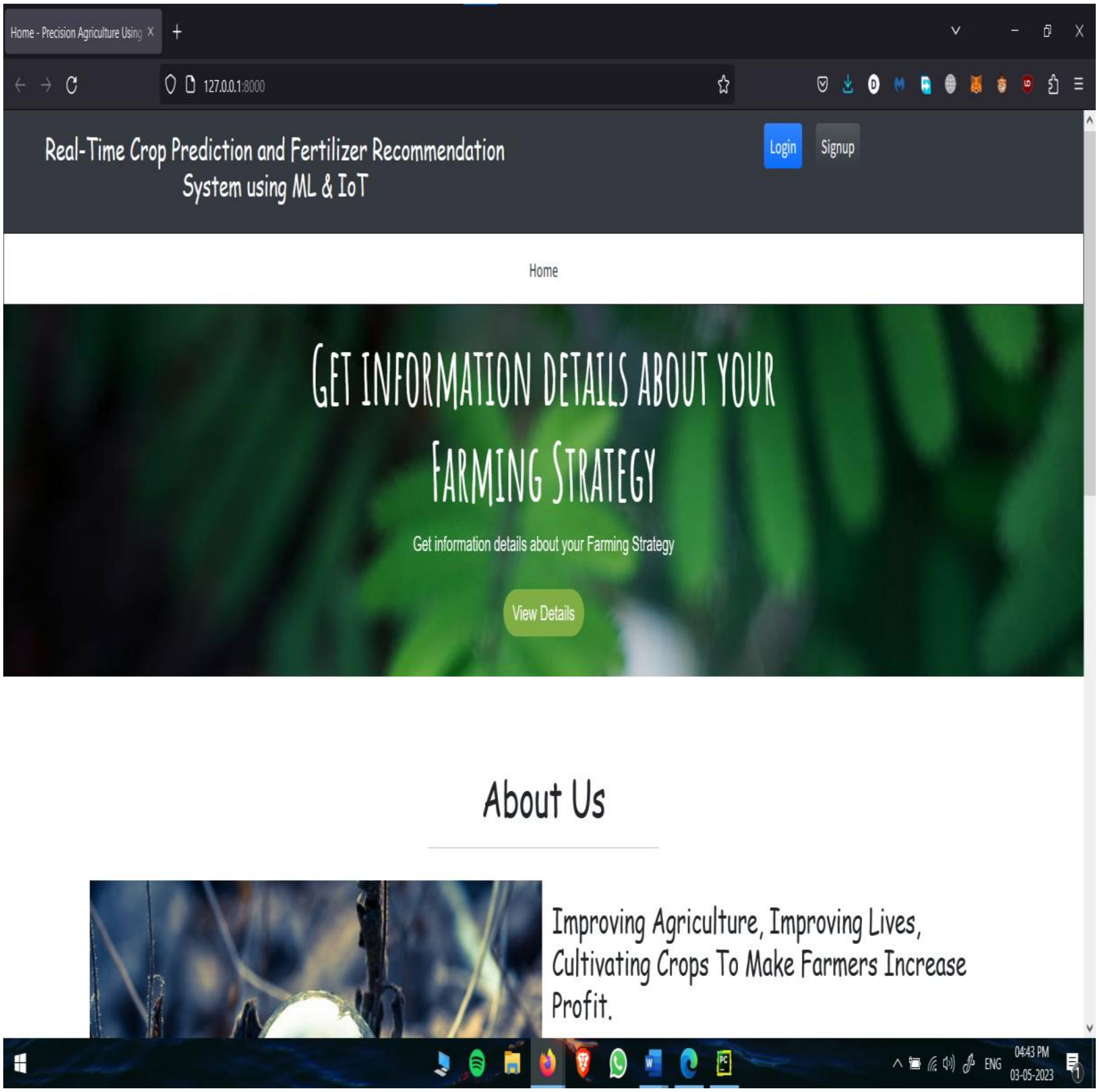
4. After successful installation of requirements, type

Python app.py runserver

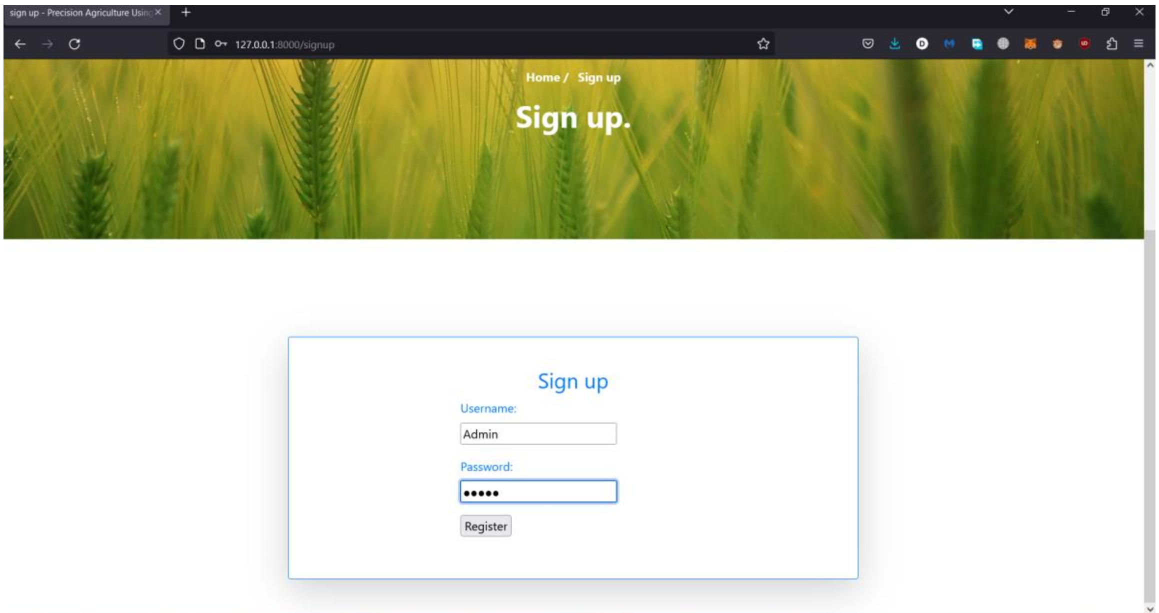
You will get following output with link to open web application on localhost



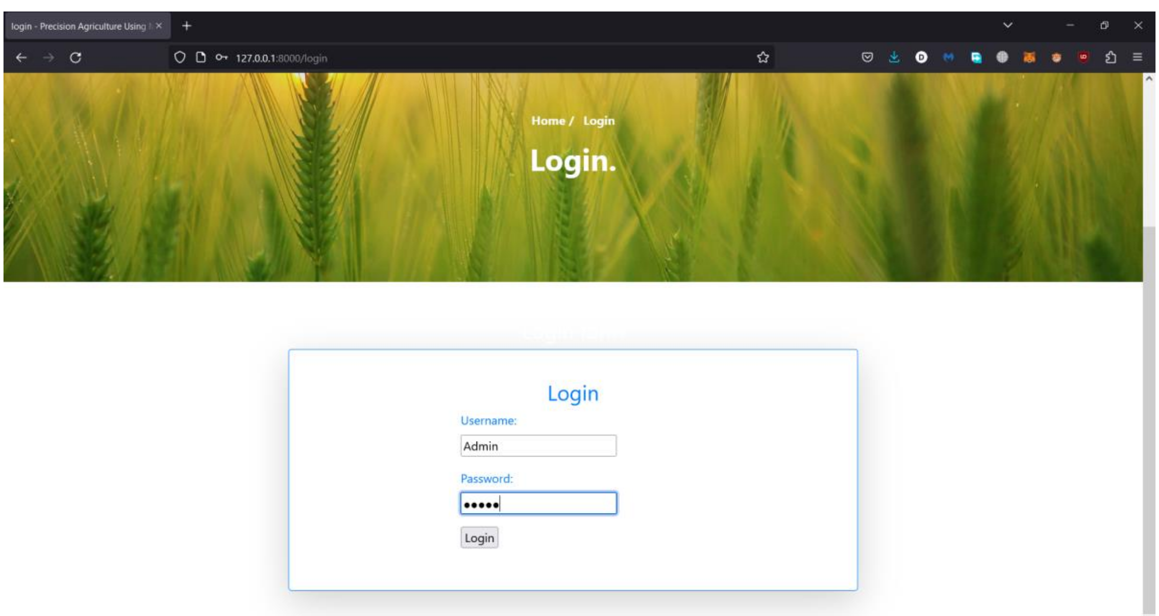
5. After clicking and opening link in web browser, you can use web application for project



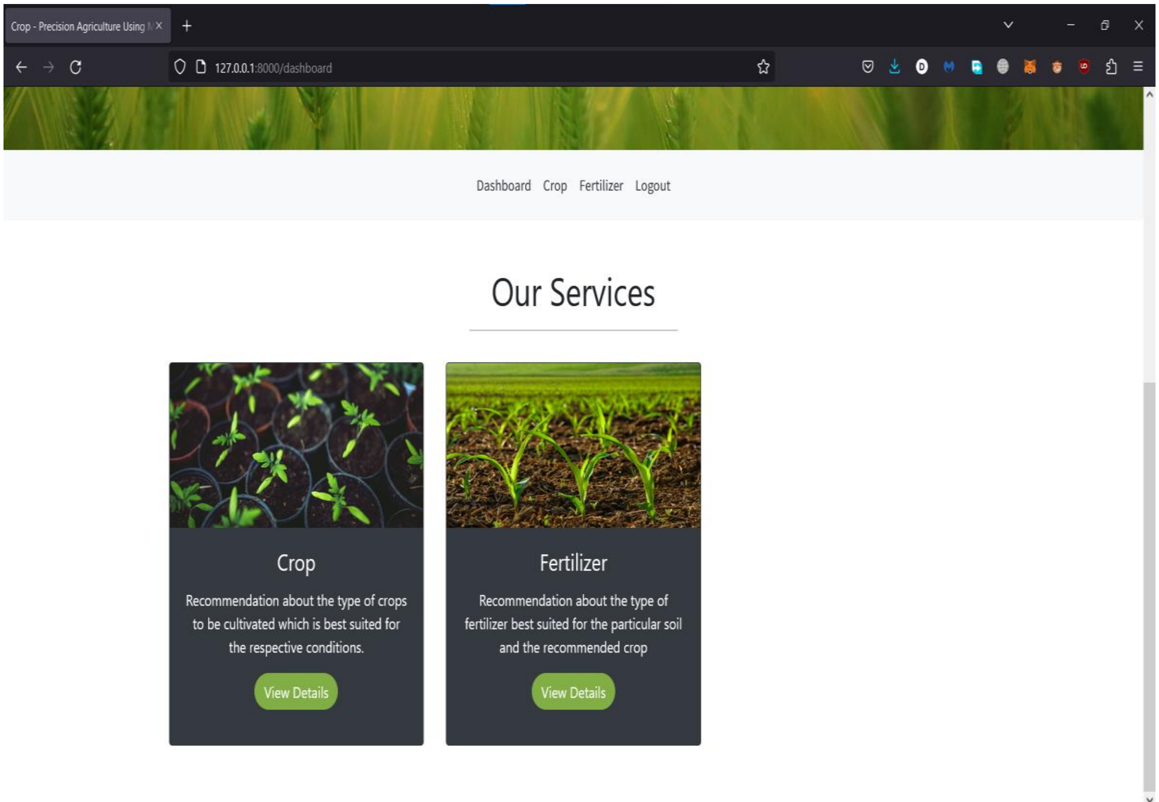
6.User can signup themselves on web app using signup button



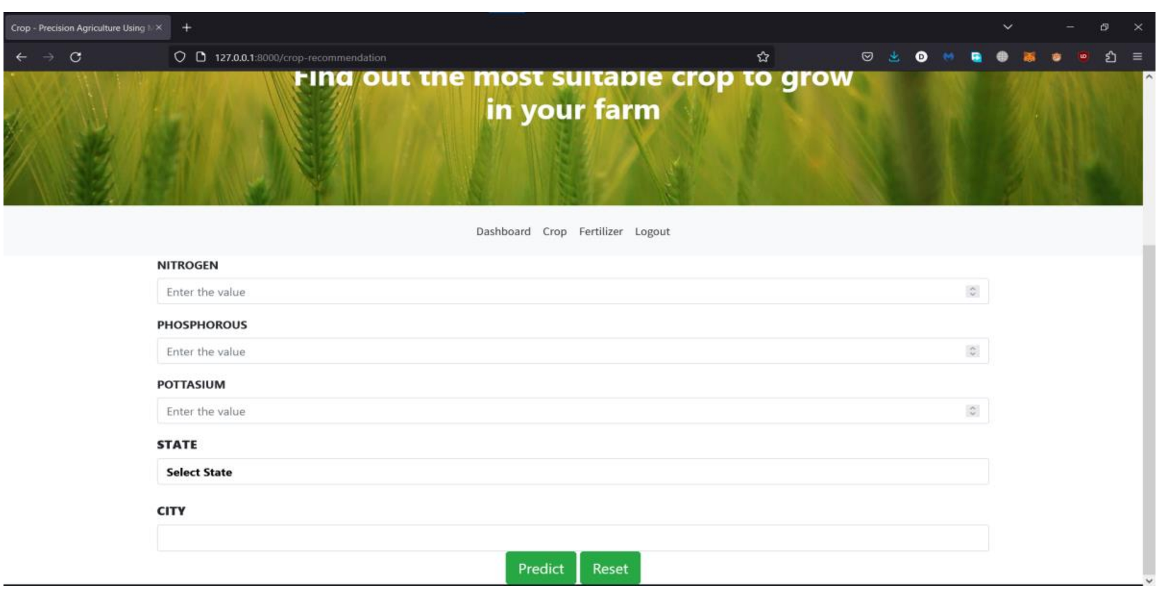
and then login to use functionalities of site.



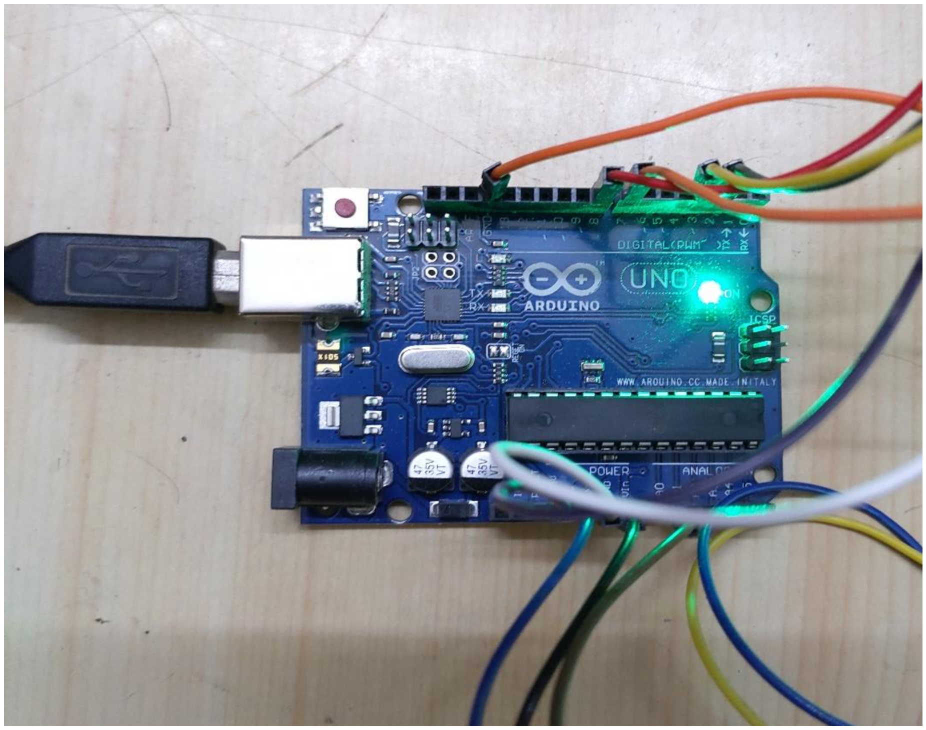
7. After login you can see dashboard like this



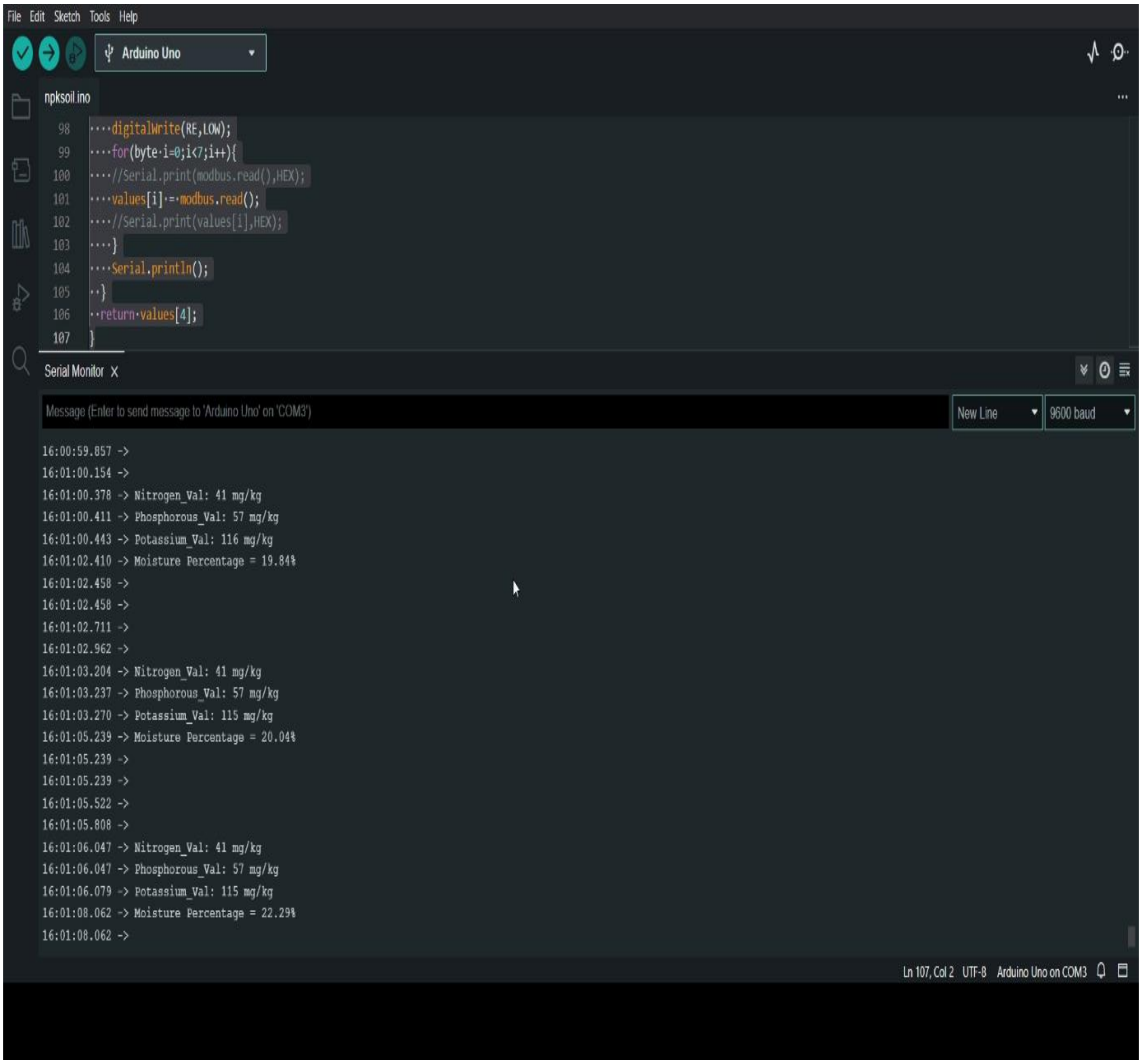
8. After selecting crop option following screen will appear



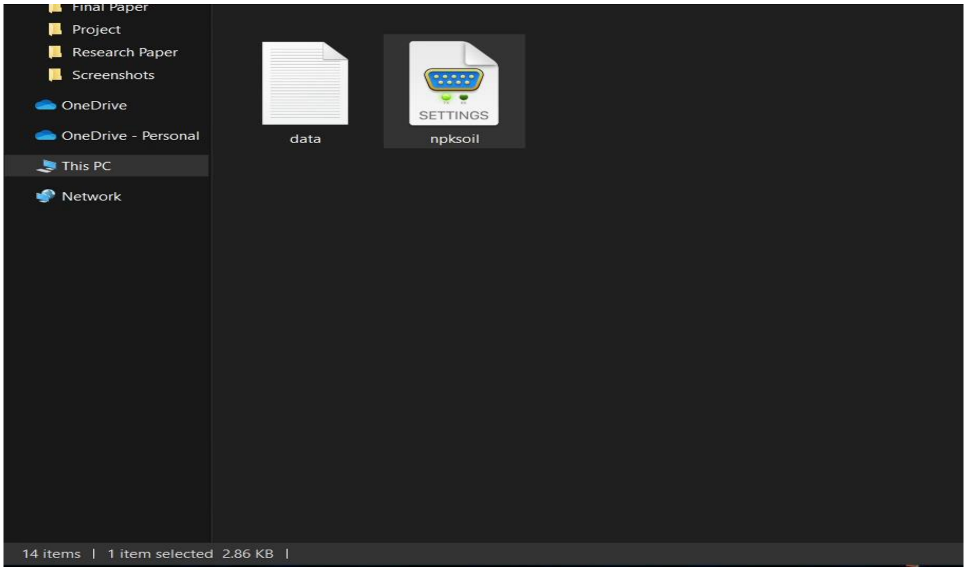
9. Now user needs to enter data for N, P, K and state, We need hardware for this step



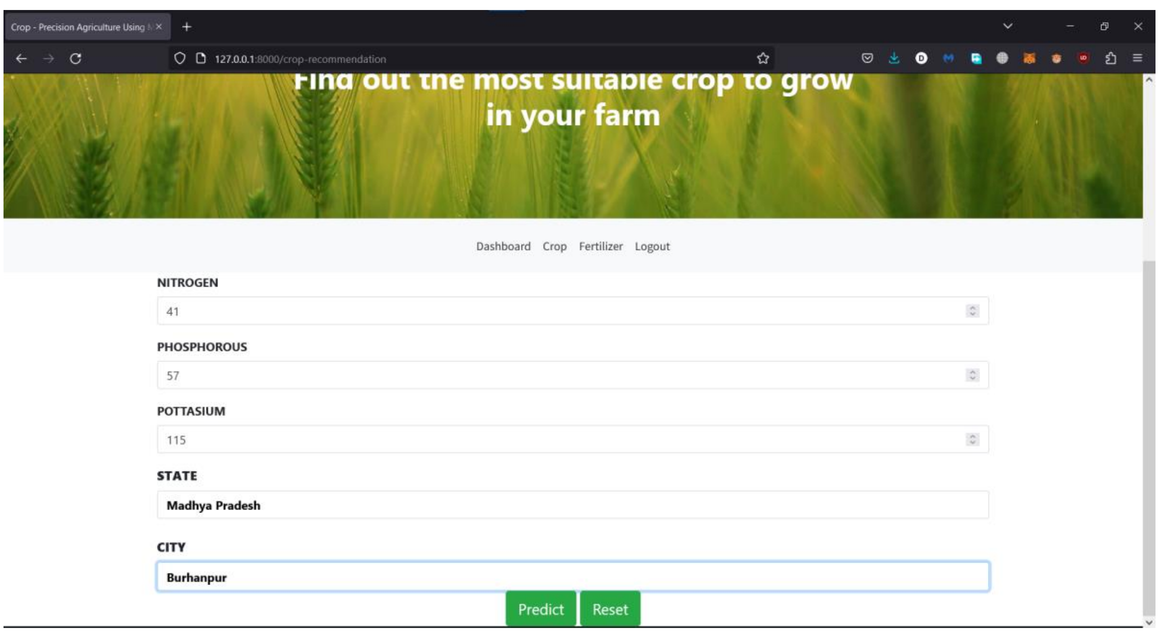
10. After connecting Arduino to PC we can get sensor into Arduino Ide as shown below



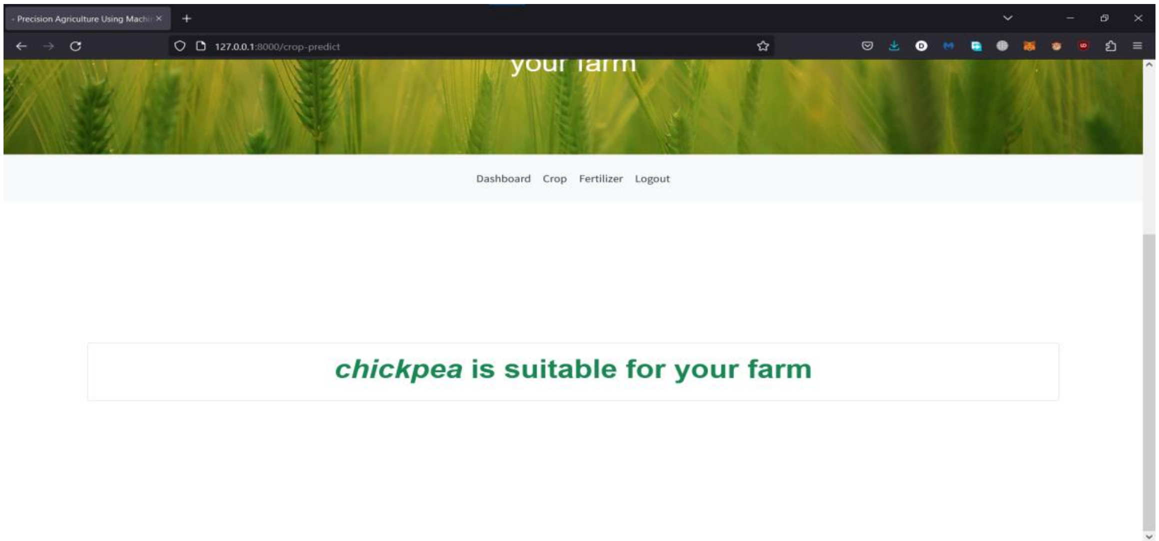
To capture this data into a text file we are using CoolTerm software which directly loads and captures Arduino code and create a data.txt file.



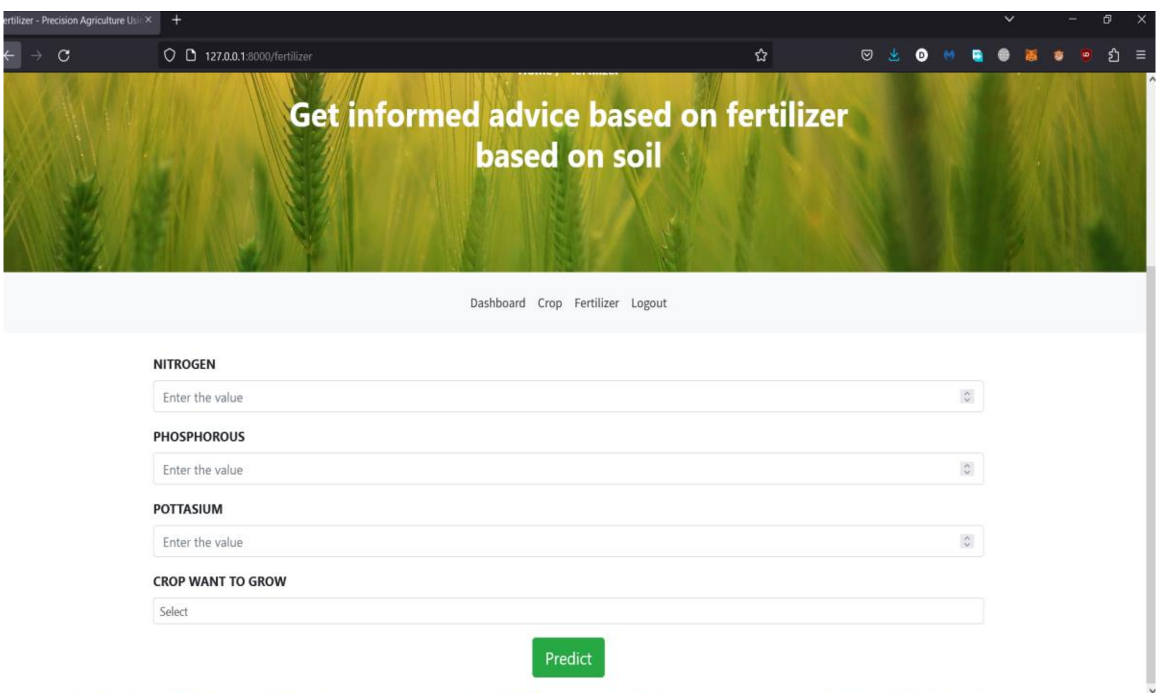
11. Now enter data collected into the web application



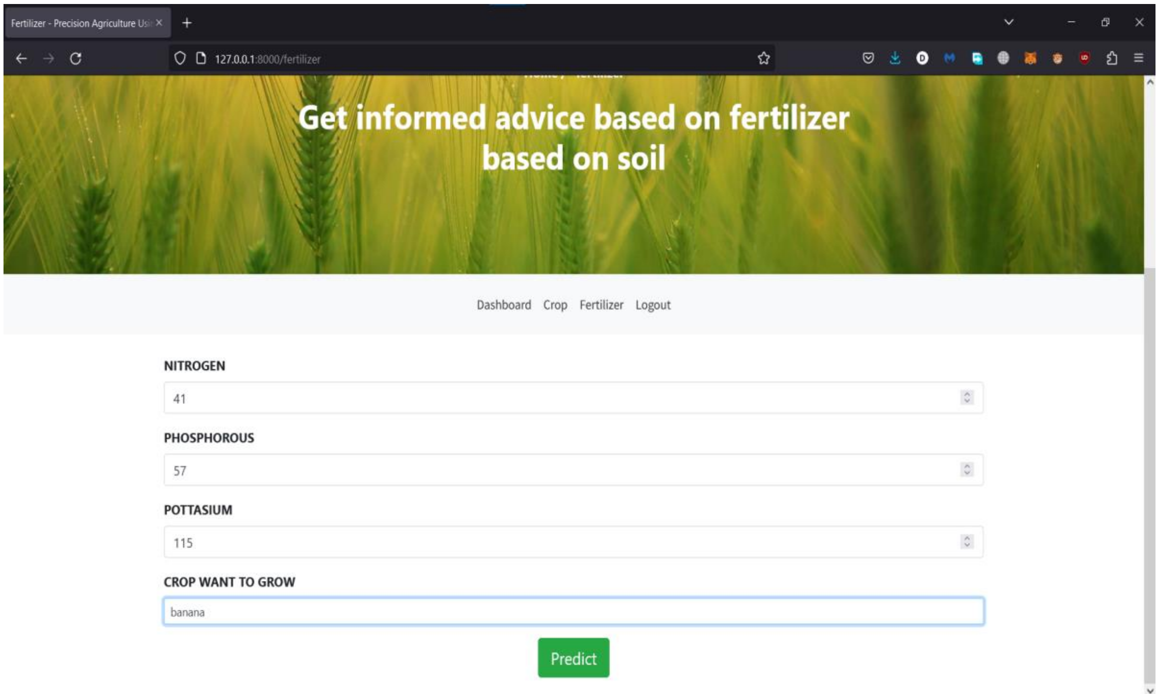
After clicking predict button result will be shown like this



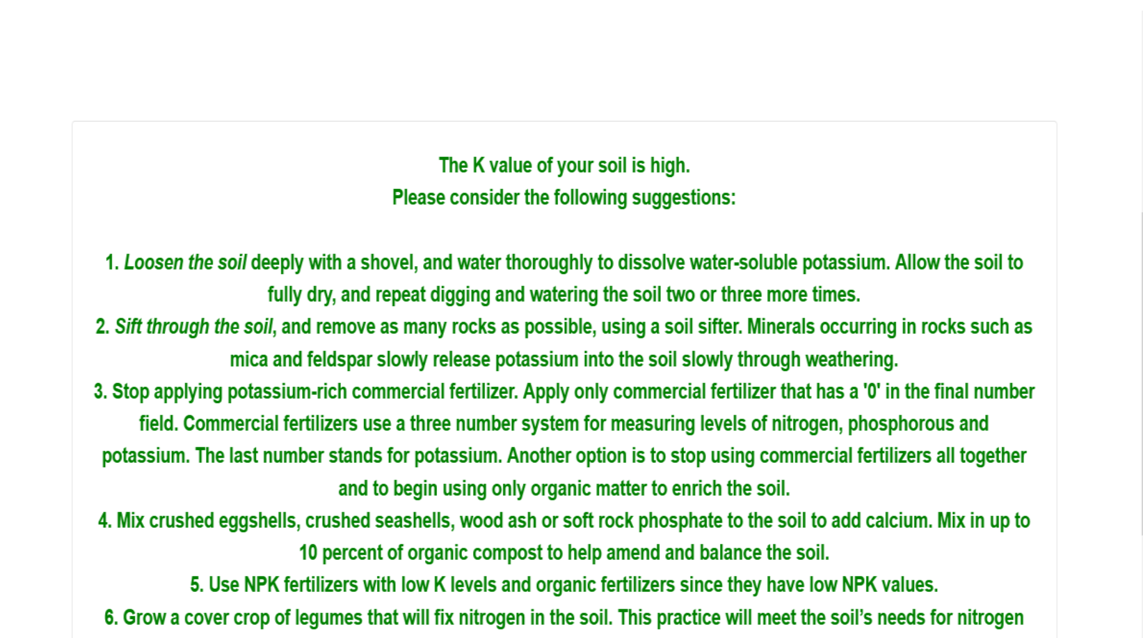
12. Now to get fertilizer recommendations similar steps need to be followed, click on fertilizer.



Input data got from sensors and crop to grow click predict.



The result will be shown like this



REFERENCES

[1] Aruvansh Nigam, Saksham Garg, Archit Agrawal "Crop Yield Prediction using ML Algorithms", International Journal of Engineering Research & Technology (IJERT), 2019.

[2] R. Kaur and S. K. Soni, "A review of crop prediction models and remote sensing-based approaches for crop yield forecasting," International Journal of Agricultural and Environmental Information Systems, vol. 8, no. 1, pp. 1-22, 2017.

[3] R. Garg, P. Jain, and R. K. Jain, "Crop yield prediction using machine learning algorithms: A review," Journal of Agricultural Science and Technology, vol. 21, no. 3, pp. 501-516, 2019.

[4] S. Saha, S. Ghosh, and M. Pal, "A review on machine learning approaches for crop yield prediction," International Journal of Computer Applications, vol. 182, no. 27, pp. 23-30, 2018.

[5] D. Kumar, A. Kumar, and S. Kumar, "Crop yield prediction using machine learning techniques: A review," Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 2, pp. 163-178, 2018.

[6] N. Dubey and D. Kumar, "A survey of machine learning techniques for crop yield prediction," International Journal of Computer Sciences and Engineering, vol. 7, no. 4, pp. 416-423, 2019.

[7] S. K. Roy, S. C. Paul, M. S. Islam, and M. A. Hasan, "Crop yield prediction using machine learning algorithms: A comprehensive review," IEEE Access, vol. 7, pp. 78907-78926, 2019.

[8] P. Patil, A. Shinde, and M. Kadam, "A survey on crop yield prediction using data mining techniques," International Journal of Advanced Research in Computer Science, vol. 8, no. 1, pp. 421-425, 2017.

[9] S. Kumar, M. R. Gupta, and M. S. Bhatia, "Crop yield prediction using machine learning algorithms: A comprehensive review," International Journal of Computer Applications, vol. 179, no. 6, pp. 1-6, 2018.

[10] M. Kamble, D. Kumar, and P. Kumar, "A review on crop yield prediction using machine learning techniques," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 7, no. 4, pp. 260-267, 2017.

[11] H. Yu and S. Liu, "A survey on crop yield prediction models and methods," Journal of Northeast Agricultural University, vol. 25, no. 1, pp. 1-10, 2018.

[12] H. Ma, X. Zhang, Y. Ma, and Y. Wang, "A survey on crop yield prediction based on deep learning," in 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), Shanghai, China, 2018, pp. 335-340.

[13] S T. K. Das and S. S. Mohanty, "Prediction of crop yield using machine learning techniques: A review," International Journal of Agricultural and Environmental Information Systems, vol. 10, no. 3, pp. 1-16, 2019.

[14] V. N. Nguyen, Q. N. Nguyen, and T. N. Huynh, "A comparative study of machine learning algorithms for crop yield prediction," in 2020 International Conference on Advanced Technologies for Communications (ATC).

[15] S. R. Singh, A. Singh, and V. P. Singh, "Crop yield prediction using machine learning: A review," Journal of Big Data, vol. 6, no. 1, pp. 1-27, 2019.

[16] D. K. Pal and B. Datta, "Crop yield prediction: A review," Journal of Agrometeorology, vol. 16, no. 2, pp. 121-130, 2014.

[17] S. Kumar, N. Kumar, and D. Kumar, "Crop yield prediction using machine learning algorithms: A comprehensive review," in 2018 IEEE 4th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 68-72.

[18] N. K. Gupta and V. G. S. Kumar, "Crop yield prediction using machine learning algorithms: A review," in 2018 International Conference on Smart Computing and Informatics (SCI), pp. 647-652.

[19] R. Dubey and S. Jain, “A review of crop yield prediction models and methods,” Journal of Engineering Science and Technology, vol. 13, no. 9, pp. 2720-2737, 2018.

[20] A. Sahay and P. N. V. N. N. Murthy, “Crop yield prediction using machine learning: A review,” in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1-6.

[21] S. G. S. Saini, S. S. Rajput, and S. S. Khillare, “Crop yield prediction using data mining techniques: A review,” in 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 1953-1957.

[22] M. Manasa, S. S. Kumar, and A. Kumar, “A review on crop yield prediction using machine learning techniques,” in 2018 International Conference on Computing, Power and Communication Technologies (GUCON), pp. 901-905.

[23] P. N. Shinde, S. K. Mahajan, and R. P. Gohad, “A review on crop yield prediction using data mining techniques,” in 2018 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), pp. 1-5.

[24] S. K. Singh, R. Singh, and S. Gupta, “A review on crop yield prediction using machine learning algorithms,” in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 433-437.

[25] T. H. Shroff and D. R. Patel, "A review on crop yield prediction using machine learning techniques," in 2019 International Conference on Intelligent Sustainable Systems (ICISS), pp. 1167-1171.

[26] A. Mukherjee, S. K. Bandyopadhyay, and T. K. Basu, "Crop yield prediction using machine learning: A review," in 2020 International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1-5.

[27] P. Singh and P. Kharya, "A review of crop yield prediction models and techniques," International Journal of Research in Engineering, Science and Management, vol. 3, no. 3, pp. 12-16, 2020.

[28] V. Pandya and S. Goyal, "A comprehensive review on crop yield prediction using machine learning," in 2021 International Conference on Inventive Systems and Control (ICISC), pp. 1757-1762.

[29] H. B. Kumar, V. Singh, and A. Kumar, "A review on crop yield prediction using machine learning techniques," in 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 114-119.

[30] S. S. Saini, S. S. Khillare, and S. S. Rajput, "A review on crop yield prediction using machine learning and data mining techniques," in 2018 2nd International Conference on Advances in Electronics, Computers and Communications (ICAECC), pp. 1-5.

[31] M. Bajgai, D. Grgicak-Mannion, and A. Mukherjee, "A comprehensive review of crop yield prediction using machine learning methods," Journal of Agricultural Informatics, vol. 12, no. 2, pp. 1-25, 2021.

[32] M. Sharma, N. Mishra, and N. Singh, "A survey on crop yield prediction using machine learning techniques," in 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 348-352.

[33] D. Liu, J. Zhang, and J. Sun, "Crop yield prediction using machine learning: A review," Journal of Physics: Conference Series, vol. 1634, no. 1, p. 012050, 2020.

[34] S. Gupta, S. Kumar, and S. S. Iyengar, "A review on crop yield prediction using machine learning," in 2018 IEEE 2nd International Conference on Inventive Systems and Control (ICISC), pp. 1006-1010.

Dissemination of Work

1. Title - Machine Learning and IoT Applications in Agriculture.

Author - Atharva Raut, Bhavy Mittal, Mayur Patel, Vedant Polshettiwar

Publisher - International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

ISSN No - 2581-9429

DOI - 10.48175/IJARSCT-9416

Issue - April 2023

2. Title - Crop prediction and fertilizer recommendation system using machine learning and IoT

Author - Atharva Raut, Bhavy Mittal, Mayur Patel, Vedant Polshettiwar, Prof. Sagar Padiya

Publisher - International Journal of Scientific Research in Engineering and Management (IJSREM)

ISSN No - 2582-3930

DOI – N/A

Issue - April 2023

Paper 1 :

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 6, April 2023

Machine Learning and IoT Applications in Agriculture

S. D. Padiya[1], Atharva Raut[2], Bhavy Mittal[3], Mayur Patel[4], Vedant Polshettiwar[5]
Assistant Professor, Department of Information Technology[1]
Student, Department of Information Technology[2,3,4,5]
Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India

Abstract: *Agriculture is an industry that has historically relied on traditional methods for crop production, but with the advent of new technologies, it is now possible to integrate machine learning and Internet of Things (IoT) applications to improve agricultural practices. Machine learning algorithms and IoT devices can be used to analyze data collected from agricultural fields to optimize crop yield, reduce resource consumption, and improve farm management. In this review paper, we explore the various applications of machine learning and IoT in agriculture, specifically focusing on their use in crop monitoring, disease detection, and water management. We examine the challenges associated with implementing these technologies in agriculture, including issues related to data collection, privacy, and security. Finally, we discuss the potential benefits of integrating machine learning and IoT in agriculture and identify future research directions that can help advance this field. Overall, this review highlights the potential of machine learning and IoT technologies to revolutionize agriculture and improve food security in the years to come.*
The Internet of Things (IoT) network must be integrated with sensors in order for "smart agriculture" to be a reality. At many layers of the IoT system architecture, machine learning (ML) techniques are incorporated to increase usefulness and capabilities. For agricultural systems to properly integrate with information technology, intelligent agricultural systems must be established, and all types of information created by agricultural systems must be integrated and analysed.The agriculture sector might undergo a transformation thanks to the fusion of machine learning (ML) and internet of things (IoT) technology. Precision agriculture and more economical resource usage are made possible by using IoT sensors to collect data on a variety of factors, including soil moisture, temperature, and nutrient levels. Then, using these data, ML algorithms may be used to forecast outcomes and improve decision-making. For example, they can forecast agricultural yields, spot disease or insect infestations, and suggest the best dates for planting and harvesting.

Keywords: Soil NPK Sensor, Soil Moisture Sensor, Random Forest

I. INTRODUCTION

Machine learning and Iot technologies have been used in agriculture to conduct innovative researches1. Using iot-based data-driven farm management strategies, agricultural yields may be improved. The key benefits of machine learning and iotin agriculture are the reduction of issues, crop and yield predictions, management of livestock, and identification of databases of agricultural output in India. By using these technologies to accurately automate agricultural cultivation practices, management, and production, researchers are assisting farmers. Agriculture is a crucial sector that plays a vital role in global food security. In recent years, there has been an increasing demand for sustainable and efficient agricultural practices that can meet the growing food demands of the world's population. Machine learning (ML) and Internet of Things (iot) are two technologies that have gained significant attention in the field of agriculture. These technologies have the potential to transform traditional farming practices by enabling farmers to make data-driven decisions, optimize crop yield, and reduce resource consumption.

In this review paper, we explore the applications of ML and IoT in agriculture, with a specific focus on their role in crop monitoring, disease detection, and water management. We discuss the challenges associated with implementing these technologies in agriculture, such as the need for large amounts of data, privacy and security concerns, and the cost of deploying and maintaining IoT devices.[1]

Paper 2:

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Crop prediction and fertilizer recommendation system using machine learning and IoT

Atharva Raut[1], Bhavy Mittal[2], Mayur Patel[3], Vedant Polshettiwar[4], Prof. Sagar Padiya[5]

[12345]*Department of Information Technology, SSGMCE, Shegaon, Maharashtra*

---***---

Abstract – Agriculture is of utmost importance for survival, and machine learning (ML) and IoT can provide valuable solutions for crop selection issues. Therefore, this study aims to predict crops suitable for farmers' soil by utilizing various ML and IoT techniques. The classification model employed is Random Forest as its showing the highest accuracy. The application of ML algorithms to predict crops and suggest which fertilizers to use can aid farmers in making decisions regarding which crops to grow based on factors such as NPK values, humidity, temperature, and rainfall thereby bridging technology and the agricultural sector.

Key Words: soil moisture sensor, npk sensor, random forest

1.INTRODUCTION

Agriculture occupies 60% of India's landmass and provides for the fundamental necessities of 1.4 billion individuals. For the well-being of the farmers, modernization of agriculture is carried out today. The production of crops is heavily reliant on weather conditions, environmental shifts, rainfall (which can often be unpredictable), water management, and the use of pesticides. This makes it difficult for farmers to achieve their desired crop. To address this issue, researchers have turned to machine learning techniques to improve crop yield and quality.[2]

Machine learning is a method of teaching machines without the need for explicit programming, it improves machine execution by portraying the consistency and pattern of information. In this work various machine learning methods such as Decision Tree, Naïve bayes, Logistic Regression, and Random Forest were employed to predict crops for different states.[7]

Since agriculture is the most dominant activity across all cultures and civilizations throughout history. It's not only a significant aspect of the growing economy but is also crucial for the survival and the future of humankind. It is a major contributor to employment and has seen an exponential increase in production requirements over time. However, people are misusing They can also be used in landscaping and gardening to conserve water and avoid over-watering. Moreover, soil moisture sensors are extensively used in research to monitor variations in soil moisture over time and analyze the correlation between soil moisture and plant growth. Soil moisture sensors operate within a voltage range of 3.3 to 5 volts.
technology to produce mass quantities, resulting in the creation of new hybrid varieties that do not provide the essential contents found in naturally produced crops, ultimately leading to environmental damage. Most of these unnatural techniques are employed to avoid losses, but with accurate information on crop, producers can minimize losses.[7] Machine learning, a rapidly growing approach, is spreading to every sector and helping make viable decisions to maximize its applications. Most gadgets today are worked with by models examined before the arrangement. The primary objective is to use machine learning models to boost agricultural sector throughput. Due to the relatively high number of parameters, the prediction is influenced by the knowledge provided during the training period. Precision agriculture, which prioritizes quality over unfavorable environmental factors, is the primary focus. Different machine learning classifiers, like logistic regression and random forest, are used to find a pattern in order to make accurate predictions and combat the erratic trends in temperature and rainfall. We came to the conclusion that the random forest algorithm provides the most accurate value by utilizing the aforementioned machine learning classifiers.[8] By collecting historical data on temperature, weather, and other factors, the system predicts which crops will be grown. The application runs the calculation and showcases the rundown of reasonable harvests for the entered information with the anticipated yield.

2. Methodology

The system relies on sensors to sense the soil parameters in real time and use an external dataset. The data obtained in real-time is stored on a local machine, and machine learning algorithms are employed for subsequent analysis.

A. Soil Moisture Sensor

A device designed to measure the moisture content of soil is known as a Soil Moisture Sensor. It is an electronic device that calculates the amount of moisture present in the soil.

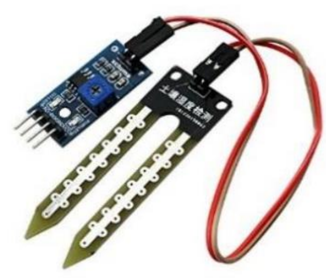


Figure 1: Soil Moisture Sensor

It typically comprises a probe that is placed in the soil to gauge its moisture content based on its electrical

Source code Listings

1. Model – RandomForest.pkl

2. Templates –

    aboutus.html

    crop.html

    crop-result.html

    dashboard.html

    fertilizer.html

    fertilizer-result.html

    index.html

    login.html

    reg.html

    signup.html

    try_again.html

3. utils –

    fertilizer.py

4. app.py

5. config.py

6.requirements.txt

7. database.db

INFORMATION OF MEMBERS

Name: Atharva Diliprao Raut

Email: atharvaraut.work@gmail.com

Mobile: 9881560104

Address: 52/1, Ratnchaya, Sangani Nagar, Ravinagar, Amravati(MS) - 444607

Name: Mayur Dipak Patel

Email: maydippat14@gmail.com

Mobile: 7507576125

Address: Mahesh Colony, Gajanan Mandir Road, Warud(MS)

Name: Bhavy Mittal

Email: bhavymittal45@gmail.com

Mobile: 9111887914

Address: 02,20 Ramganj ward budhwara bazar khandwa MP

Name: Vedant Shashikant Polshettiwar

Email: vedantpolshettiwar10@gmail.com

Mobile: 8975789412

Address: Behind Ramakrishna Hotel, Zulelal colony, Wardha (MS)