Blockchain and Data Analytics-based Crop Recommendation System for Precision Farming in IoT Environment

By

Devangi Patel
20BCM012

Guided By

Dr. Rajesh Gupta
[DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING]

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Ahmedabad 382481

NIRMA UNIVERSITY
INSTITUTE OF TECHNOLOGY
NAAC ACCREDITED 'A+' GRADE

CERTIFICATE

This is to certify that the Computer Engineering Project entitled "Blockchain and Data Analytics-based Crop Recommendation System for Precision Farming in IoT Environment" submitted by Devangi Patel (20BCM012), towards the partial fulfillment of the requirements for the degree of Integrated B.Tech.(CSE)-MBA of Nirma University is the record of work carried out by him/her under my supervision and guidance. In my opinion, the submitted work has reached the level required for being accepted for examination..

Dr. Rajesh Gupta
Assistant Professor
Computer Science and Engineering Dept.,
Institute of Technology,
Nirma University,
Ahmedabad

Dr. Madhuri Bhavsar,
Professor and HOD,
Computer Science and Engineering Dept.,
Institute of Technology,
Nirma University,
Ahmedabad

ACKNOWLEDGEMENT

We would like to express my heartfelt gratitude and appreciation to our Head of Department, Dr. Madhuri Bhavsar, and our Course Coordinator, Mr. Chandan Trivedi for their contributions to the completion of our project titled.

In addition, I extend my sincerest thanks to our dedicated mentor, Dr. Rajesh Gupta, whose expertise and guidance have been instrumental in shaping our project. Their invaluable advice, constant availability, and commitment to our growth have been truly inspiring. He provided us with the necessary resources, share their knowledge, and ensure that we stayed on the right track.

ABSTRACT

In agriculture, soil is a vital element that decides the quality and yield of agricultural produce. Soil consists of various nutrients such as Nitrogen (N), Phosphorous (P), Potassium (k), pH, and water content. Considering soil health would provide farmers with the best selection of crops compatible with their farm's soil nutrients. In the first module of this project, we have proposed an algorithm for recommending a set of suitable crops based on various soil attributes. These soil nutrients can be collected in real-time using soil sensors such as N, P, K sensor, pH sensor, and humidity sensors, they can be deployed in farms where the cultivation takes place. The crop recommendation model will use data from these sensors in real-time which would increase the accuracy of the result. In the second module, we have proposed a solution for farmers that would benefit them with an optimized companion crop selection. Our mixed-cropping selection model employs ML algorithms to predict the compatibility among crops. The set of crops suggested by the recommender system is used as an input to the mixed-cropping selection model. The model will then calculate the compatibility score of the crops based on various attributes such as root system, water requirement, harvest period, soil type, growth rate, height, etc., and recommend the most suited crop combination.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Table 1: Abbreviations

ML	Machine learning
IoT	Internet of things
ICT	Information and Communications Technology
MQTT	Message Query Telemetry Transport
KNN	Nearest Neighbors Algorithm
SVM	Support vector machine
RF	Random forest
GB	Gradient boosting
N	Nitrogen
P	Phosphorours
k	Potassium
pH	Potential of Hydrogen

NOTATION

English Symbols

Z*	Non-negative integers
P_d	Time frame
S_n	No. of sensors
C_n	No. of Crops
N_n	Value of each attribute

CHAPTER 1

INTRODUCTION

1.1 Introduction

For the last many decades, our entire ecosystem is highly dependent on agriculture. Agriculture involves the scientific process of cultivating soil and growing crops. It includes practicing various activities such as horticulture, aquaculture, livestock production, and farming. Farming is not only essential to a country's economy but also to individuals, as the food we consume comes from farms. Globally, India is entitled as the second most populous country which leads to a greater demand for food in India. Thus, to enhance crop yield and increase agricultural outputs, farmers employ a variety of methods such as crop rotation, use of organic fertilizers, adopting modern irrigation techniques, and intercropping. In India, intercropping was traditionally done using a method called "mixed cropping," where two or more crops were grown together in the same field. This practice allows farmers to maximize the use of available resources and reduces the risk of crop failure. This raised a concern as farmers usually pick a crop based on their instincts without considering salient factors such as soil health, weather conditions, and market price. The consequences for the same include bad harvest and less profit for farmers.

The share of the labor force working in the agriculture sector has seen a substantial downfall in over a decade. The rationale behind this leads to multiple challenges faced by farmers. But there is one which directly impacts the overall productivity and quality of crop yield. It is Soil degradation, in simple terms it means that the soil will lose its original quality which in turn affects its fertility and formation. It results in poor soil health, thus reducing the capacity of agricultural produces that soil can render. There are various ecological factors responsible for soil degradation but in this paper, we will focus on the influence of human activity and its impact on soil health.

Inconvenient agricultural practices such as monoculture planting, overgrazing of animals, row cropping, tilling or plowing, crop removal, and land-use conversion can lead to soil degradation. The most commonly practiced among these is monoculture planting. It is a technique of cultivating the same crop in large fields year after year. The significant issues faced due to monoculture planting can be Nutrient deficiency, Nutrient imbalance, and susceptibility to pesticides. Every crop has its unique requirements and nutrient intake from the soil, if the same crop is planted it will frequently consume a specific set of nutrients which can lead to nutrient deficiency in the soil. Even while practicing row cropping if crops having the same nutrient requirements are grown together, they will fight for nutrients and eventually will result in poor yield and depletion of the crop quality, thus leading to nutrient imbalance. Monoculture builds an environment where a crop is targeted with specific pest immunity. To prevent pests, farmers will have to use heavy pesticides which can bring changes in soil formation which will make these pesticides ineffective over a period of time, thus leading to diminished soil productivity and soil degradation.

1.2 Need of project

Farmers presume that adding readily available fertilizers can improve the quality of their yield in a short period. Repeated use of these fertilizers can lead to the production of contaminated harvests and reduce the long-term fertility of the soil. Thus, it is important to understand specific nutrient deficiencies in the soil and look for an appropriate solution.

This invites the need for a fertilizer recommendation system that suggests, fertilizers based on a nutrient deficiency in the soil. This results in optimized use of resources and leads to a better crop yield.

To practice mixed cropping, farmers needed a system that would provide them with the best crop combinations. The mixed-cropping selection model employs ML algorithms to predict the compatibility among crops. The set of crops suggested by the recommender system is used as an input to the mixed-cropping selection model. The model will then calculate the compatibility score of the crops based on various attributes such as root system, water requirement, harvest period, soil type, growth rate, height, etc, and recommend the most suited crop combination.

1.3 Organization

This report is a composition of five chapters illustrating, details and working of every stage of the project.

1.3.1 Chapter 1 Summary

This chapter talks about the problems faced by farmers and current trends practised in farming. It discusses the need for the implementation of the project to solve the problems faced by the farmers.

1.3.2 Chapter 2 Summary

This chapter enlists various research papers highlighting the problems and their respective solutions. Various solutions have been proposed such as escalating the yield of farmers by considering various parameters such as soil nutrients, rainfall, temperature etc. Also, blockchain technology has been introduced for the secured transmission of sensor data. researchers have also stated that mixed cropping practices can enhance yield. Along with this the pros and cons of the research papers have been mentioned.

1.3.3 Chapter 3 Summary

This chapter presents the framework which illustrates the various stages of the recommendation algorithm and the mixed cropping ML-based model. The recommendation algorithm consists of five stages: data collection, migration, pre-processing, blockchain layer and crop recommendation model. The mixed cropping ML-based model consists of a machine learning algorithm that suggests the most suitable combination of crops that can be grown on the farm.

1.3.4 Chapter 4 Summary

This chapter shows the experimental results of the algorithms which were applied in the project. It consists of all the informative graphs and plots of the recommendation system model and mixed cropping ML-based model.

1.3.5 Chapter 5 Summary

This chapter concludes the project and states all the problems considered and their respective proposed solutions. This chapter also talks about the future enhancements of the project.

CHAPTER 2

Literature Review

2.1 Literature Review

Many researchers have given various solutions to the problems mentioned above. For example, Gupta et al. [1] presented the use of IoT-based smart farming that improved the agriculture system by monitoring the field in real-time. Factors like humidity, temperature, etc were considered and machine learning (ML) algorithms were applied which recommended a suitable crop. Further, Bhuyan et al.[2] provided a statistical look at the features and indicated the best crop type on the given features. ML algorithms like k-NN, SVM, RF, and GB trees were examined for crop-type prediction. Panigrahi et al. [3]developed an ML model to predict farm production.The data was collected and trained using supervised machine learning with six distinct regression models to estimate the crop yield. The productivity of a crop also depends on the nutrient content of the soil. However, these solutions have not considered important attributes such as nitrogen (N), phosphorus (P), Potassium (K), and pH which plays a vital role in determining the effective crop yield.

Inline with the above-said issues, Ahmed et al.[4] explored the correct usage of soil nutrients such as N, P, and K. Their exploration was used to develop a knowledge-based system using IoT sensors. Furthermore, Ahmed et al. used an improved genetic algorithm (IGA) to recommend an optimal setting for nutrients for different crops. Then, Paul et al. [5] classified soil into low, medium, and high categories using data mining techniques such as naive Bayes to predict the crop yield. Despite the various benefits of the previous approaches for crop nutrient prediction, it has some disadvantages, such as, instead of using a dataset that has values in a discrete format such as (low, mild, bad, etc), working on attributes having continuous values can lead to precise outcomes. Algorithms using nutrients like NPK can make inaccurate predictions, thus utilizing all essential attributes would lead to comparatively better crop prediction.

Soil nutrient deficiency can be treated by adding fertilizers to the soil. Regardless, farmers presume that adding readily available fertilizers can improve the quality of their yield in a short period of time. Repeated use of these fertilizers can lead to the production of contaminated harvest and reduce the long-term fertility of the soil. Thus, it is important to understand specific nutrient deficiencies in the soil and look for an appropriate solution. This invites the need for a fertilizer recommendation system that suggests, fertilizers based on nutrient deficiency in the soil. This results in optimized use of resources and leads to a better crop yield.

Information sharing in an IoT-based system through an open wireless channel can pose various security risks. Since the system is connected to the internet, it is vulnerable to cyber attacks such as hacking, data breaches, and malware. The real-time data from the sensors can be hindered, manipulated, and altered and hence the anomaly in the prediction of a crop. Since farmers rely on the app, they may grow the crop, recommended by the erroneous prediction model and therefore they are at a loss as they don't get a healthy harvest.

Cloud computing is another excellent technology that enables us to store valuable data from our sensors on the cloud, and even allows us to access it remotely. This paper on an IoT framework for precision agriculture by Bakthavatchalam et al.[6] proposes a system for interfacing real-world data using sensors. These Sensors are deployed in various geographical locations and combined to form a network. Responses acquired by these sensors then can be stored and accessed via the cloud. However, using cloud can have certain challenges such as data security and reliability. Data transmitted over the cloud can be tampered with and used for illegitimate purposes. Nevertheless, like any challenge, there is a solution to this particular problem.

Blockchain can be implemented to overcome these problems. It ensures enhanced security, increased transparency, improved traceability, faster transactions, and improved data-sharing techniques. It can be used to build a trusted and open agriculture ecosystem even when the parties in the system may not trust each other. To overcome the various security risks in the open wireless channel, various research papers have implanted Blockchain technology to ensure security in the IoT-based system. Lin et al.[7] discussed the integration of Blockchain with the IoT system highlighting the benefits and challenges. They proposed a model ICT (Information and Communications Technology) e-agriculture system with a blockchain infrastructure for use at the local and regional scale. They presented a security framework that blends blockchain technology with IoT devices. Further, Rahman et al.[8] proposed a scalable and distributed data-sharing system integrating access control for smart agriculture. It used anonymous identities to ensure users' privacy. Their proposal ensured trust, privacy, scalability, high throughput, information-intensive farming, and accessibility. Chaganti et al.[9] proposed the security monitoring framework prototype for smart farms using Arduino Sensor Kit, AWS cloud, and the smart contract on the Ethereum Rinkeby Test Network and evaluated network latency to monitor and respond to security events. They also proposed a smart farmer community-based monitoring solution.

Many researchers have studied and suggested the advantages of a mixed cropping system over monoculture plantations. As mentioned in the paper by Prieto et al.[10], mixed cropping has produced 15% higher yields than mono-crops, concerning planting multi-species crops. Another paper by Li et al. [11] says in addition to yield, there are multiple other advantages of mixed-cropping such as enhanced soil fertility by intercropping with nitrogen fixers. They also explored the effect of mixed cropping on the usage of pesticides, according to research it was inferred there is an increased resilience against pests and disease. This led to increasing demand for growing two or more crops simultaneously on the same piece of land, otherwise known as Mixed - cropping. Mixed cropping not only combated all the issues faced by farmers while practicing monoculture but also reduced the risk of crop failure.

As farmers started to practice mixed cropping, researchers saw the need to build recommendation systems that can assist farmers in selecting a set of crops that are compatible with the nutrient content present in their croplands. In the paper by Pudumalar et al.[12], a crop recommendation system is constructed based on various soil attributes such as soil pH, water holding capacity, soil texture, etc. this system will suggest crop using various ML algorithms, to farmers that meet the soil requirements and will increase the productivity thereby increasing the profitability. on precision farming, where researchers have implemented a recommendation system using data analytics. the algorithm used in this paper suggests multiple crops based on their accord with various soil nutrients such as n,p,k, and other attributes of soil health.

Despite the benefits provided by the recommendation models, there is one limitation noticed in all. You see, suggesting a set of crops is not sufficient, engaging compatible crops is imperative to guarantee optimum advantages of practicing mixed cropping. As mentioned in the paper by Nasaret et al.[13] crops grown in a mixed cropping system tend to interact with each other, positively and negatively. Incompatible crops may require similar nutrients which could lead to resource competition amongst them. Using similar kinds of crops could also lead to issues like pest susceptibility, as crops with common vulnerabilities will make them more susceptible to insects and pests. Thus having insightful crop combinations can lead to better results and improved crop yield using available resources.

Thus we have proposed a solution for farmers that would benefit them with an optimized companion crop selection. Our mixed-cropping selection model employs ML algorithms to predict the compatibility among crops. The set of crops suggested by the recommender system is used as an input to the mixed-cropping selection model. The model will then calculate the compatibility score of the crops based on various attributes such as root system, water requirement, harvest period, soil type, growth rate, height, etc, and recommend the most suited crop combination.

2.2 Summary

Chapter 2 enlists various research papers which highlight the problems and their respective solutions. Various solutions have been proposed such as escalating the yield of farmers by considering various parameters such as soil nutrients, rainfall, temperature etc. also, blockchain technology had been introduced for secured transmission of sensor data. researchers have also stated that mixed cropping practices can enhance the yield. along with this the pros and cons of the research papers have been mentioned.

Table 2.1: Comparative study of research papers

Author	Year	Objective	Methodology	Pros	Cons
Gupta *et al.* [1]	2021	Smart crop prediction using IoT	Decision Tree, KNN, and Support Vector Machine (SVM)	presented the use of IoT-based smart farming in order to predict suitable crop	Soil Nutrients such as nitrogen, phosphorous, and potassium are not considered
Bhuyan *et al.* [2]	2023	Crop type prediction using machine learning approach	k-NN, SVM, RF, and GB	proposed work aims to statistically study the attributes' role and impact on crops' growth also indicated the best crop type on the various weather and soil features	
Panigrahi *et al.* [3]	2023	Crop Yield prediction using supervised learning	Linear Regression, Decision Tree Regression, Gradient Boosting Regression, Random Forest Regression, Xgboost Regression, And Voting Regression	analyzed how machine learning techniques can be employed to develop an accurate crop yield prediction model	researchers were concentrating more on crop yield rather than crop selection
Ahmed *et al.* [4]	2021	Nutrient recommendation system	IGA(improved genetic algorithm)	proposed an improved genetic algorithm (IGA) to recommend an optimal setting for nutrients to different crops	Not considered attributes such as soil pH, temperature, humidity
Paul *et al.* [5]	2015	Soil behavior and crop yield prediction using data mining	Naive bayes	classified soil into low, medium, and high categories using data mining techniques to predict the crop yield	working on a dataset having discrete format such as (low, mild, bad, etc), which could lead to less accurate results
Bakthavatchalam *et al.* [6]	2022	Iot framework for measurement and precision agriculture	Multilayer perceptron, JRip, Decision table	smart module in recommending the crop for irrigation and obtaining maximum yield based on present environmental factors.	Sensor data is stored on the cloud/backend, which can lead to potential security breach
Lin *et al.* [7]	2018	Integration of blockchain and iot	ICT e-agriculture model	It presented a security framework that blended the blockchain technology with IoT devices.	It did not consider the scalability of the blockchain system
Rahman *et al.* [8]	2020	Distributed data-sharing system integrating access control for smart agriculture	used anonymous identities to ensure users' privacy	proposal ensured trust, privacy, scalability, high throughput, information intensive farming and accessibility among the users.	It did not say much about the consensus and cryprographic algorithms and interoperability
Chaganti *et al.* [9]	2022	Blockchain-based security monitoring using IoT in agriculture	used Arduino Sensor Kit, AWS cloud , and the smart contract on the Ethereum Rinkeby Test Network	enabled remote monitoring,security alerts and negligible network latency	It did not consider the scalability of the blockchain system

CHAPTER 3

Proposed Framework

3.1 Proposed Framework

The presented framework demonstrates the working of various stages of the crop recommendation system. The system consists of six stages namely, data collection, data migration, data preprocessing, use of blockchain technology, implementation of crop recommendation algorithm, and user dashboard. Each stage is elucidated as follows:

3.1.1 Data Collection

The dataset we created for crop recommendation model was a result of the compilation of raw data from multiple online references. The dataset covers five main attributes namely Nitrogen(N), Phosphorous (P), Potassium(k), pH, and Water content. In addition, it is composed of 49 different agriculture produces which when classified botanically consist of crops, fruits, and vegetables. The values in the dataset indicate a

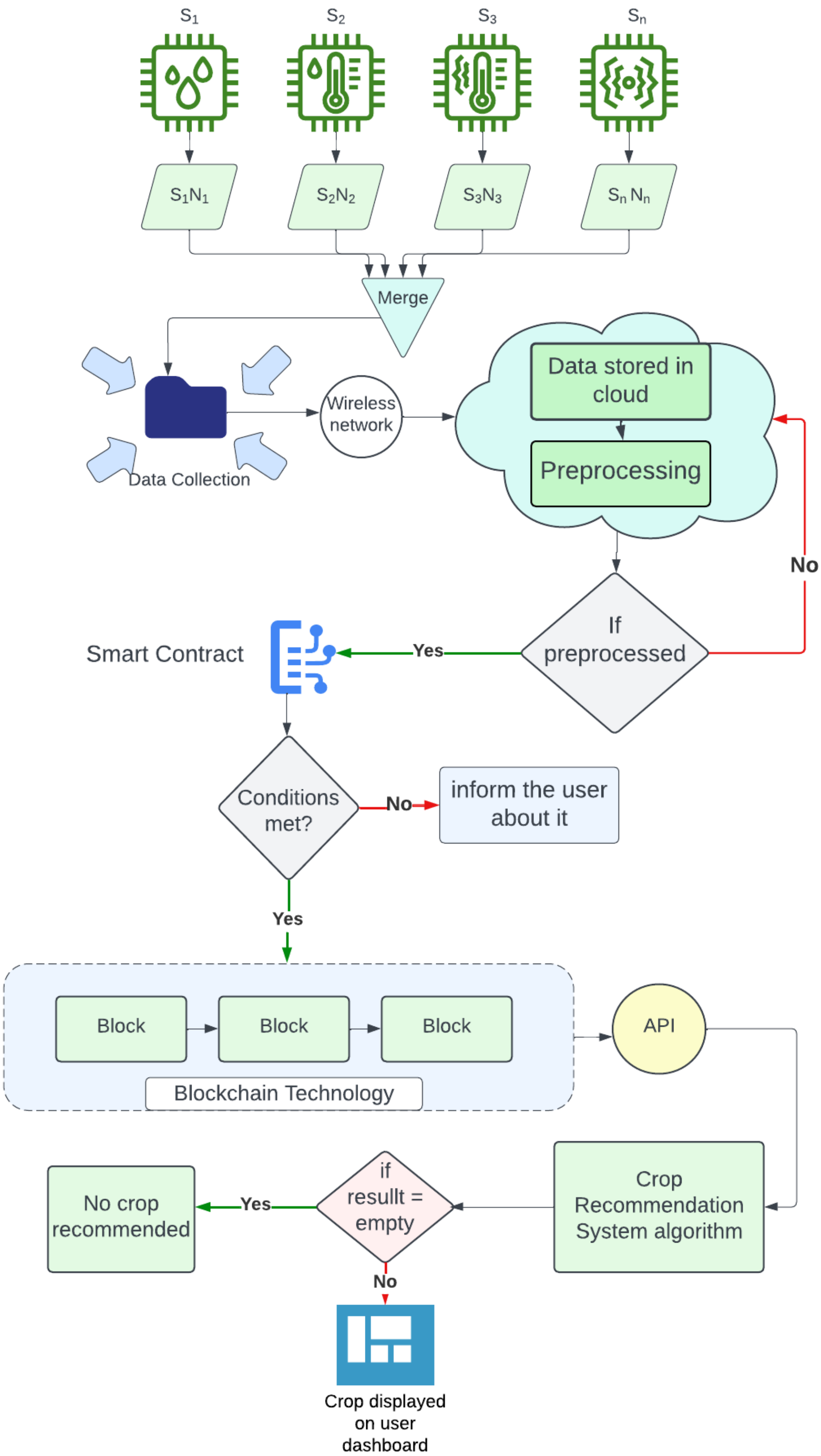


Fig. 3.1: Working of Crop recommendation system

desirable range of (N, P, K) content, pH, and water content in the soil corresponding to every crop. In our first stage, we install sensors on the farms to collect data on relevant soil attributes such as soil pH, Nitrogen, Phosphorous, Potassium, and water content. Data can be derived from an observation or a measurement of any event thus sensors such as soil nutrient sensor, pH sensor, and humidity sensor provides us with real-time data of soil constituents and conditions. Consider N$\{x|x = 5$ and attributes measured$\}$, and SN (reading of S different sensors) where SN = $\{x|x \in$ to Z*$\}$. S_iN_1 is reading for first attribute lets say, nitrogen for sensor(S_i). These sensors will be programmed to collect the data in CSV format, which then will be further migrated for processing.Table3.1 shows the sample data from N sensors for time frame P_t. P_tS_i represents the sensor data from sensor i at a time frame i where i goes from 1 to n and t goes from 0 to d.

Table 3.1: A sample data from n sensors for time frame(P_d)

Time	Nitrogen(S1)	Potassium(S2)	Phosphorous(S3)	pH(S4)
P_1	2.64	1.56	3.4	5.5
P_2	2.856	1.62	3.65	5.6
P_3	2.94	1.6	3.428	5.5
P_4	3	1.59	3.6	5.7

3.1.2 Data Migration/Transmission

After the data has been collected, the next step is to transmit the data from the sensors to the cloud server. The data generated from the sensors is sent to the cloud with the help of cellular network. Cellular network works well with large coverage area. Cellular modems are used as the medium to modulate the analog signal generated by the sensors so that the data can be transmitted using cellular networks. During the migration of the data to the cloud server, various network protocols such as Message Query Telemetry Transport (MQTT) etc. is also used. The data is then stored at the cloud server from where it can be fetched whenever needed.

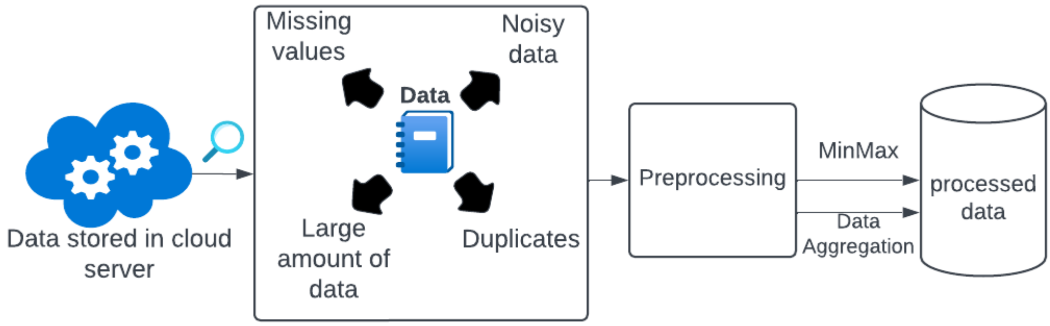


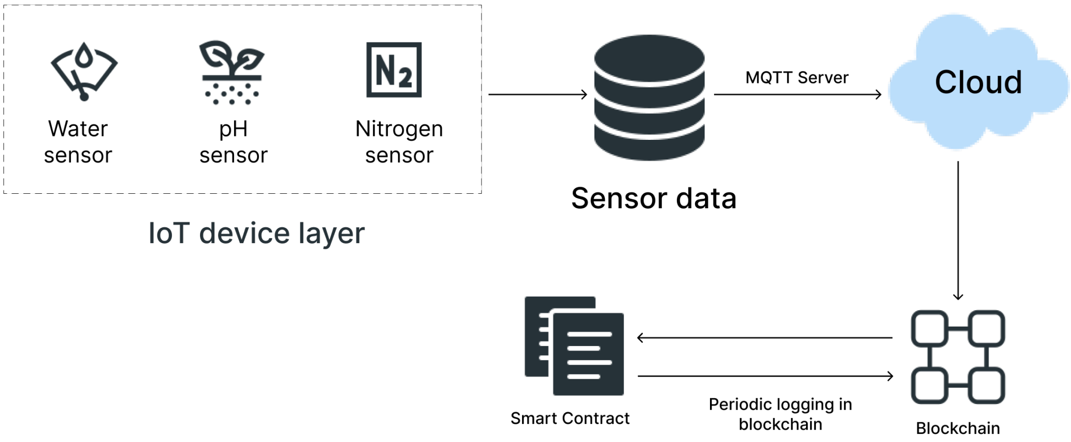
Fig. 3.2: Preprocessing steps.

3.1.3 Data Prepossessing

Data preprocessing is required to eliminate inconsistencies from the data. Inconsistencies may include missing values, errors, noise, duplicate values, data belonging to different units and ranges. This makes the data difficult to read and hence needs to be processed. The sensors may be damaged or malfunctioned or damaged due to which sensors may give missing values. These missing values can be removed by interpolation, extrapolation, smoothing or with the help of machine learning algorithms. We removed the missing values by interpolation technique where the value of the missing values was determined with the help of nearby valid data.

The average of the valid data was taken and missing values were replaced by that value. Moreover, the data can include values in different ranges and units and thus the data needs to be normalized so that we do net get anomalies while using this data for the recommendation of a crop. Normalization is required so that the values which belong to different ranges and units are normalized between 0-1. Min-Max algorithm was used to implement the normalization in our data from the sensors. Also, there is large amount of data generated from the sensors every second which in turns results in large amount of data. It would be easier if the data is aggregated and stored which would lead to efficient data analysis.

3.1.4 Blockchain Layer



The fourth stage involves use of blockchain technology. Clean and processed sensor data received from cloud, is moved to next layer. This layer would validate sensor reading i.e n,p,k, pH readings from sensors, and ensure they are tamper-proof and evident. Each block on the blockchain will store these reading from IoT devices along with its timestamp, thus maintaining a record of real-time data fetched from sensors. Smart contract automates the process and eliminates the need for dependency on third parties. Thus, no other entity has to verify that the data coming from sensor is from a genuine actor. Storage contract has several attributes such as shown in fig3.1.6. It includes farmer's name, address, crop name, batch id, soil nutrient values and timestamp. Upon satisfying all the conditions in smart contacts the data is then stored in blockchain. This monitors the authenticity of all the sensor reading as no one would be able to delete or modify data stored in this blockchain.

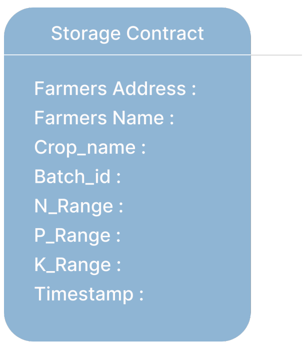
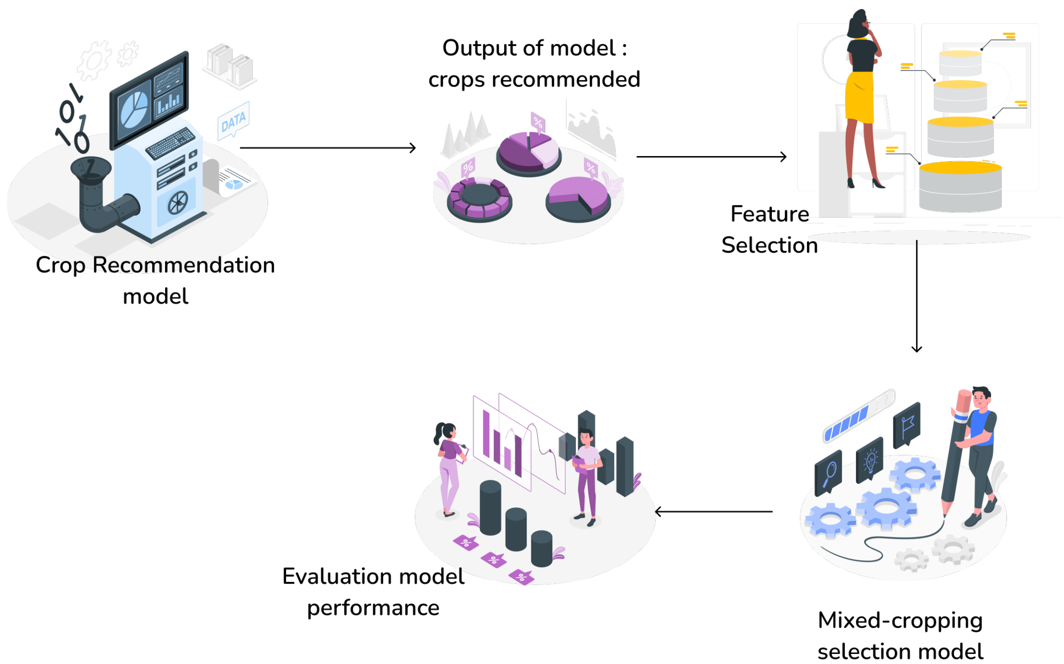


Fig. 3.3: Smart contract attributes

3.1.5 Crop Recommendation Model

Consider, we have n sensors $\{S_1, S_2, S_3, \ldots, S_n\} \in S$ used for sensing the nutrient content of the soil. Consider n soil nutrients $\{S_1N_1, S_2N_2, S_3N_3, \ldots S_nN_n\} \in S_iN_i$ corresponding to reading of each S_i sensor for given time frame. Also, consider a dataset with crops $\{C_1, C_2, C_3, \ldots, \ldots C_r\} \in C$ where r is the total number of crops which corresponds with $\{C_1N_1, C_2N_2, C_3N_3, \ldots, C_nN_n\}$ as a range of recommended values corresponding to those N attributes. The algorithm checks if SN_i a subset of C_iN_i $(\{i \in \{1, \ldots, n\})$. If it is a subset then, C_y is the recommended crop. If n-1 readings of S_iN_i are not the subset of C_iN_i, then the algorithm will recommend increasing the S_iN_i such that it becomes a subset of S_iN_i. In Algorithm **??** (line 1), we created an array R. Then(line 2-5), the loop starts and compares the sensor reading, S_iN_i with Lower limit(L) and upper limit(U) of C_nN_n. If SR[i] satisfies the condition, then (line 6) we append C[i] Corresponding to array R[]

3.1.6 Mixed-cropping model



This model takes the crops suggested by the recommendor system as the input. After the feature selection is executed, the data is passed into the ML algorithm. Random forest regressor is used to predict the compatibility of the crops. The output is analyzed and the best combination of the crops is suggested by the model.

3.1.7 User Dashboard

After implementing the crop recommendation model on sensor data, we obtained a set of the most preferred crops, suitable to various measures such as soil nutrients, pH, and water content. Growing these crops would benefit farmers, as it will not only improve their overall yield production but also allow them to do so by maximizing the use of available resources. The last stage leads us to display these results via user dashboard. The dashboard will simplify the process of monitoring and analyzing data coming from various resources on the farm such as nutrient sensor, pH sensor, and humidity sensor. It can help farmers to keep in check with farms from remote locations.

3.2 Summary

This section presented the framework which illustrateed the various stages of the recommendation algorithm and the mixed cropping ML-based model. The recommendation algorithm consists of five stages- data collection, data migration, data pre-processing, blockchain layer and crop recommendation model. The mixed cropping ML-based model consists of a machine learning algorithm which suggests the most suitable combination of crops that can be grown on the farm.

CHAPTER 4

Experimental Setup

4.0.1 Recommendation-based Results

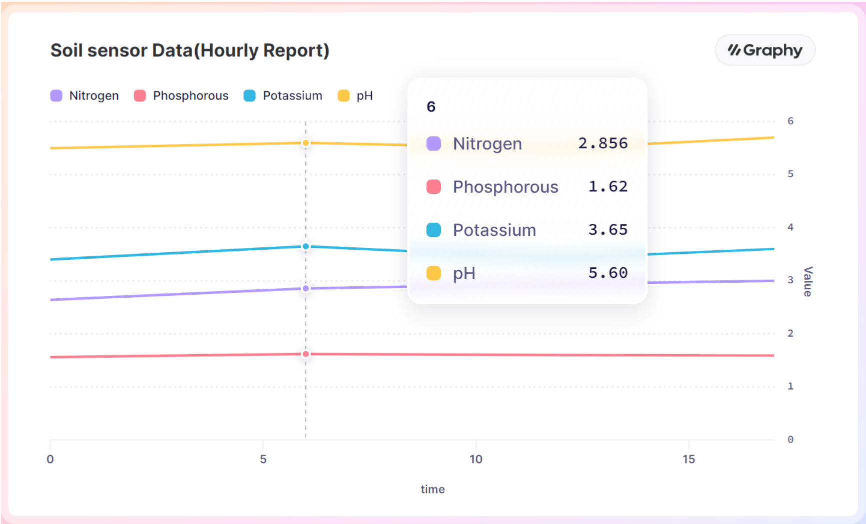


Fig. 4.1: Real time sensor data

Figure 4.1 shows the real time data from N, P, K and pH sensors at different time instances. The graph shows as the time passes the soil nutrient content remains the same.

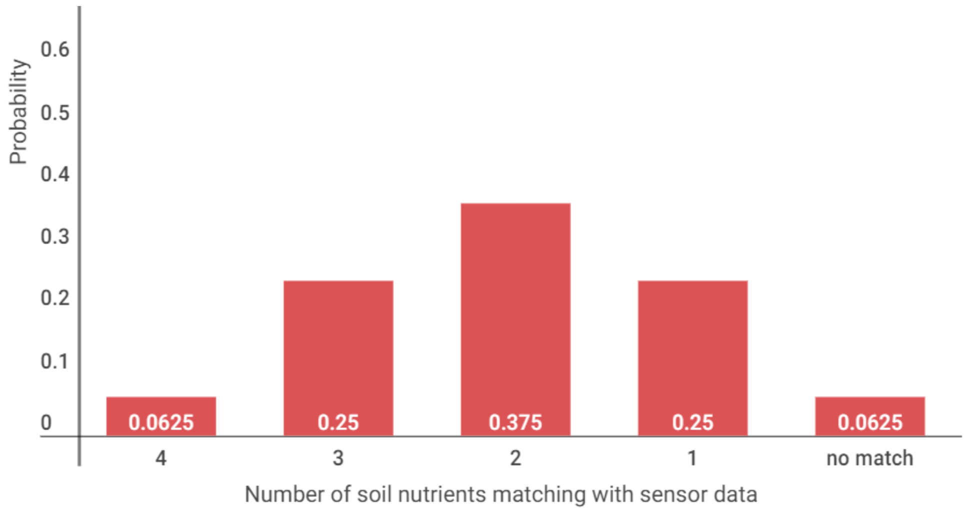


Fig. 4.2: Probability of number of soil nutrients matching the sensor data.

Figure 4.2 shows the probability of number of soil nutrients which are received from the soil sensors that fall into the recommended range of the soil nutrients of a particular crop. As can be interpreted from the graph, the probability that all four nutrients match the required range of a particular crop is the lowest. A crop will be recommended only if any of the four, three or two of the total soil nutrients fall into the required range for growing that particular crop. Otherwise, in case of only one match or no match at all, the crop won't be recommended.

Figure 4.3 depicts the favorable range of soil nutrient values of the crop mango and the corresponding sensor readings. The orange bar in the graph represents the required range of values of soil nutrients in order to grow mango. While the black line in the graph is discrete sensor values which are retrieved from the sensors deployed on a farm. All the black lines are inside the orange bar of all the soil nutrients of mango which implies that mango can be grown since all the soil nutrient requirements are satisfied.

Figure 4.4 depicts the favorable range of soil nutrient values of the crop carrot and the corresponding sensor readings. The orange bar in the graph represents the required range of values of soil nutrients in order to grow carrot. While the black line in the graph is discrete sensor values which are retrieved from the sensors deployed on a farm. Sensor readings of N, P, pH satisfy the soil nutrient requirements since they are inside the orange bar whereas the senor reading of K is not inside the orange bar which means the sensor

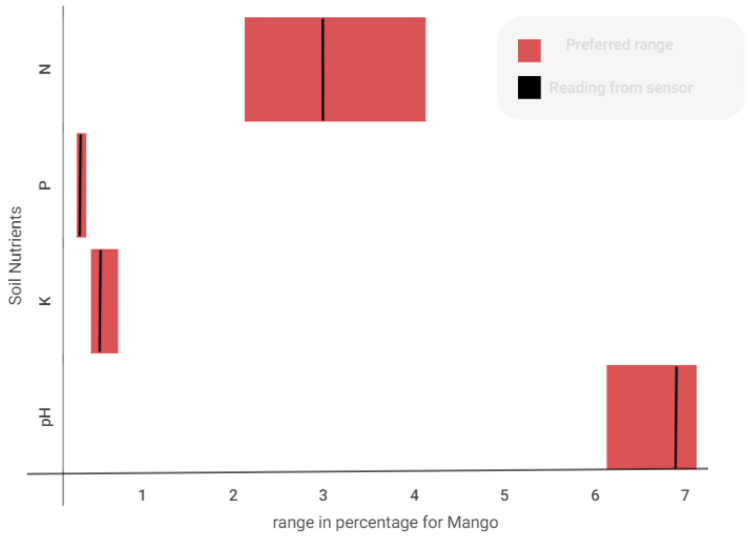


Fig. 4.3: Sensor data with 4 matching nutrients

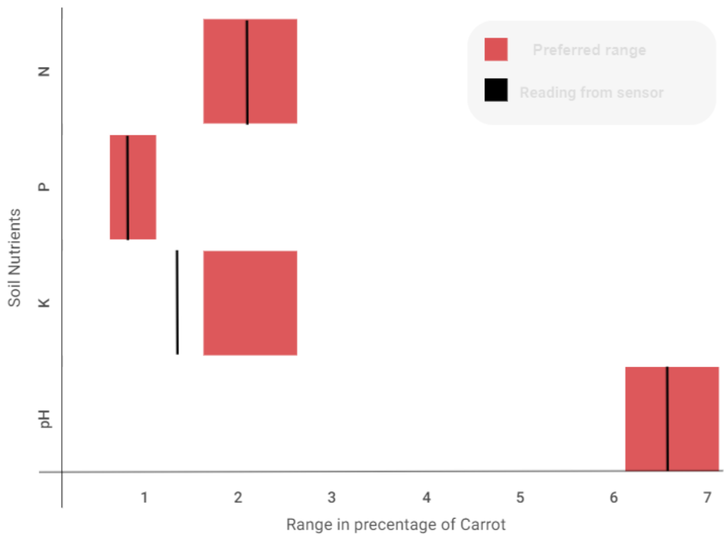


Fig. 4.4: Sensor data Graph with 3 matching nutrients

reading of K doesn't fall into the required range for growing carrot.

Figure 4.5 depicts the favorable range of soil nutrient values of the crop watermelon and the corresponding sensor readings. The orange bar in the graph represents the required range of values of soil nutrients in order to grow watermelon. While the black line in the graph is discrete sensor values which are retrieved from the sensors deployed on a farm. Sensor readings of K, P satisfy the soil nutrient requirements since they are inside the orange bar whereas the senor readings of N and pH is not inside the orange bar which

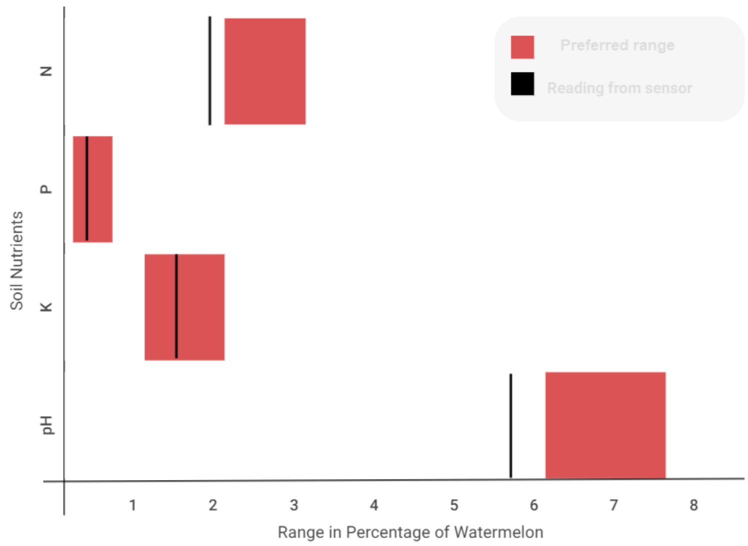


Fig. 4.5: Sensor data Graph with 2 matching nutrients

means the sensor reading of K and pH do not fall into the required range for growing watermelon.

4.0.2 Implementation interface

Figure 4.6 displays the application dashboard of the user. The dashboard tracks the various parameters of the farm and then visually depicts the same. It includes displaying total number of sensors deployed in farm from which we are receiving data, analysis of the sensor data, graphical representation of the soil nutrients, water content of the soil, and the recommended crops from the formerly mentioned parameters.

Information about the sensors

One of the features of the user dashboard is displaying the number of sensors online. It concludes the number of sensors from which we are retrieving the data and using that data in order to analyze it for the recommendation of a particular crop. As can be inferred from 4.6, there are 4 sensors online from which we are receiving the data. The weekly readings from the N, P, K sensors are portrayed graphically. The dashboard also displays the pH of the soil which can be seen as 5.4 in the 4.6.

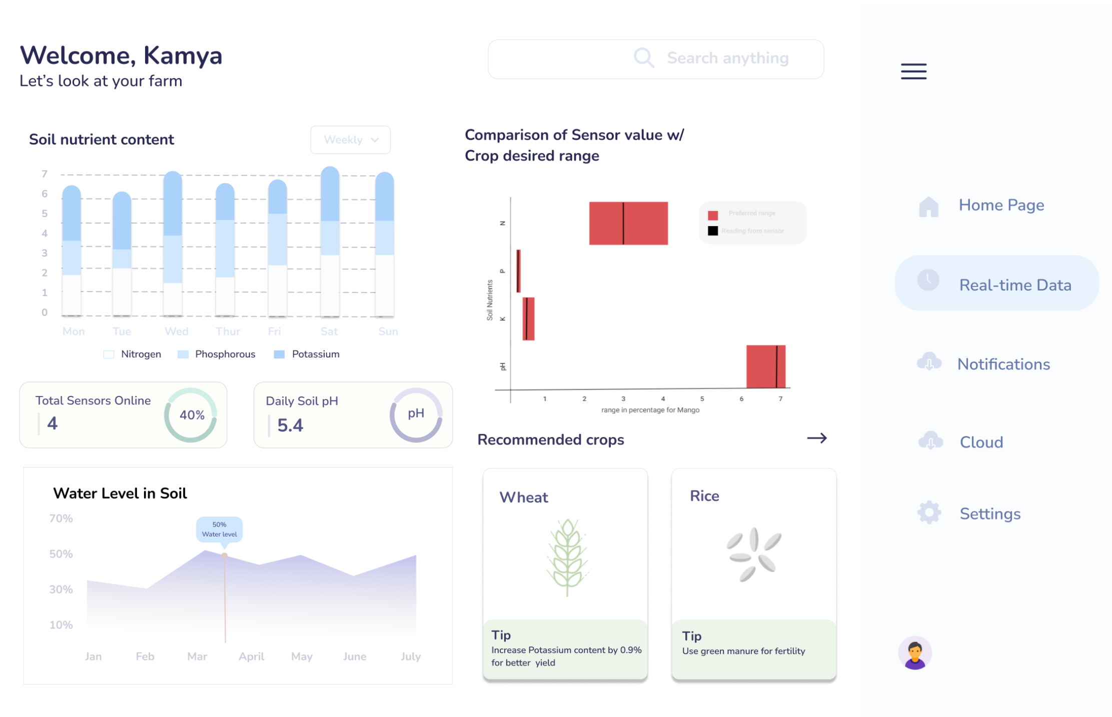


Fig. 4.6: Implementation interface.

Water content of the soil

The amount of water the soil holds is an important parameter for the healthy growth of a crop. The dashboard represents the water level in soil graphically. The graphs inputs the monthly readings of the water content of the soil. As can be inferred from the graph, the water content was the highest in the month of March.

Comparing the sensor and the desired range of soil nutrients for growing a particular crop

The dashboard also includes the graphical representation of the analysis made through the sensor data and from the dataset.

Recommended crops

At the end, after considering all the important factors, recommendation system recommends the crop which are suitable to be grown on the user's farm. As can be seen in the 4.6, wheat and rice are are recommended to the user along with the useful tips such as increasing or decreasing a particular soil nutrient content in the farm in order to ensure the healthy growth of that particular crop. The tips help to improve the production and promise a good harvest for the user.

4.0.3 Mixed Cropping ML-based Results

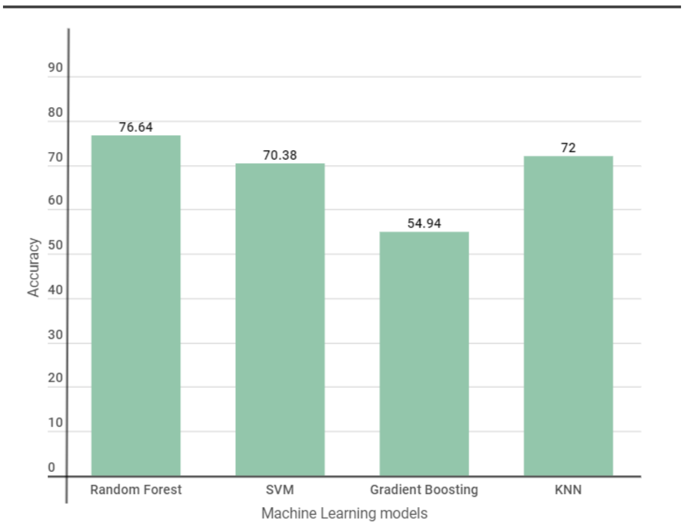


Fig. 4.7: Accuracy of the ML models

In the Figure 4.7, bar chart represents a comparative analysis of accuracy seen in different ML models. Accuracy is used to measure the performance of the different models with the intention of helping researchers in making informed decisions while selecting among distinct models Each bar represents models implemented i.e Random forest, SVM, Gradient Boosting, and KNN. From this chart, we can infer that accuracy of the Random forest model is the highest thus it is selected.

The ML model analyses the pattern and relation between the attributes and then accordingly gives the prediction and heatmaps provide a visual aid to display the correlations. Figure 4.8 is a heatmap which is the graphical representation in a matrix form that illustrates the correlation between any two attributes of the data frame. The graph finds the correlation among the following attributes: crop encoded, root system encoded, growth rate, Pest Management, water requirement, soil encoded, yield, market price, root system encoded, height, Nitrogen fixer, Soil Stabilizer, pollinator attractor, prevent soil erosion, promote biodiversity, drought conditions, waterlogging, Crop2, Compatible, percentage. It follows a colour encoding scheme where the dark colour implies that the two attributes are positively related whereas the decrease in the intensity

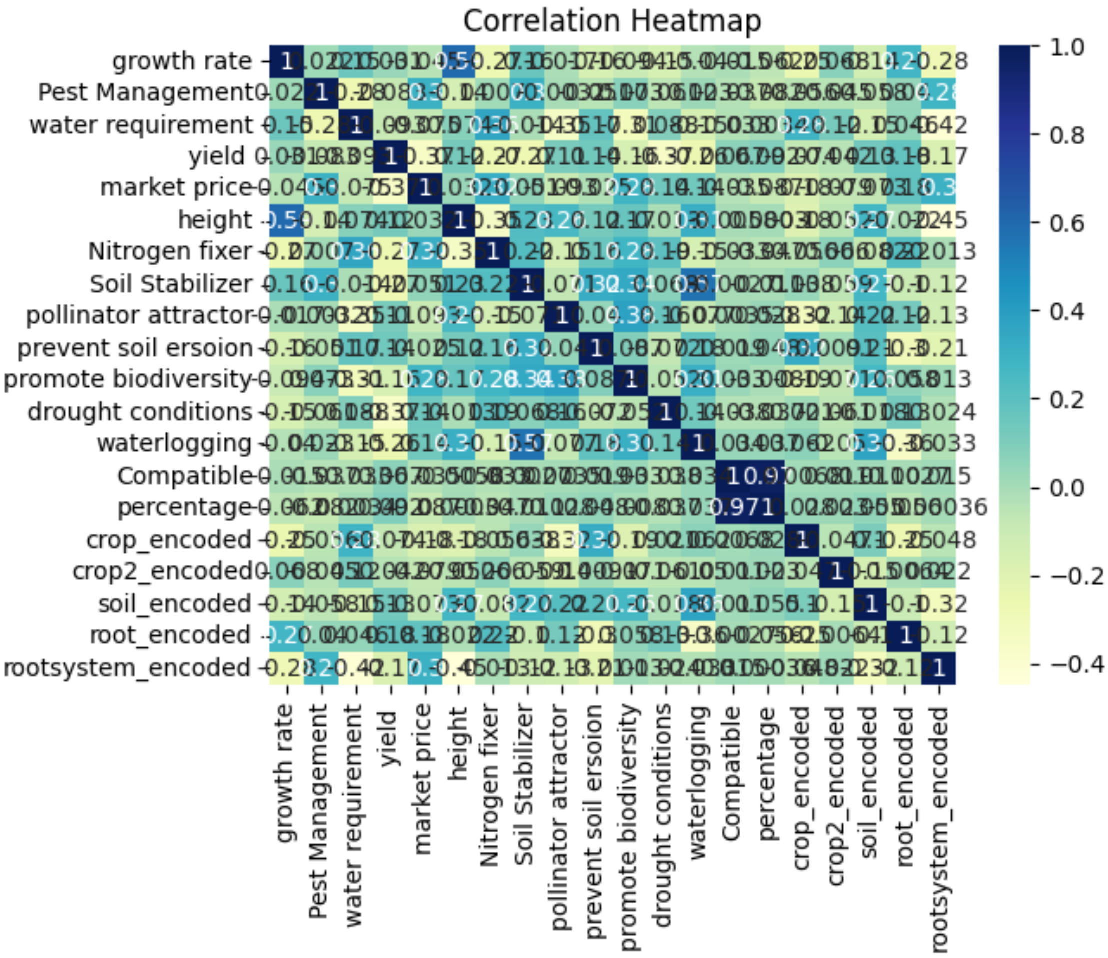


Fig. 4.8: Heatmap of the ML algorithm

of the colour signifies the opposite. Each cell in the matrix represents the correlation between the two variables. As can be inferred from the graph, the cell representing height and the waterlogging attributes has good positive relation since it is colour encoded in a shade of blue. Whereas water requirement and growth rate are negatively correlated because their cell is filled with the shade of yellow colour.

The Figure 4.9 is a residual plot. A residual plot is used to find the efficiency of the fit of an ML model graphically with the help of residuals, which are the difference between the expected and the predicted values of a target attribute. The randomness of the residuals imply that the model is capturing the pattern in the data. The graph illustrates the behaviour of the predicted values with the max difference of 0.8 from the expected values which in turn also aids in finding outliers. The points with the highest difference can be tagged as outliers which further can be removed in order to increase the accuracy of the model.

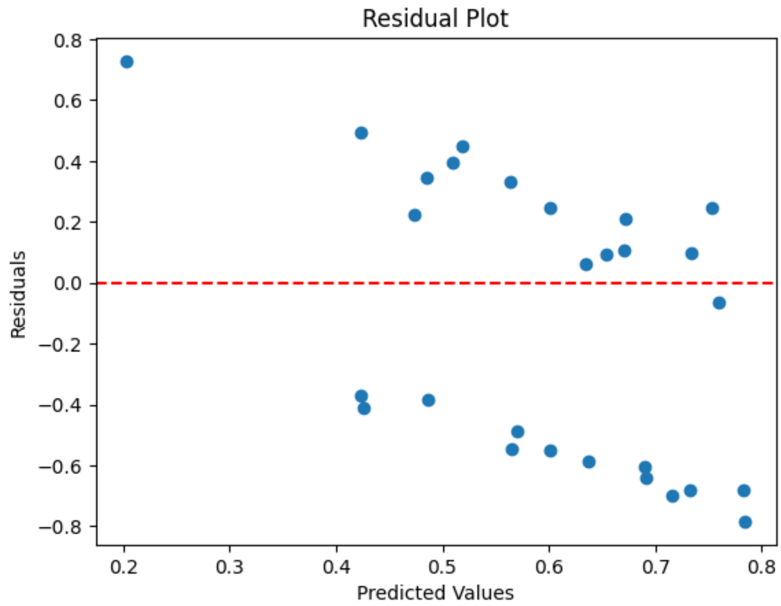


Fig. 4.9: Residual Plot of the ML algorithm.

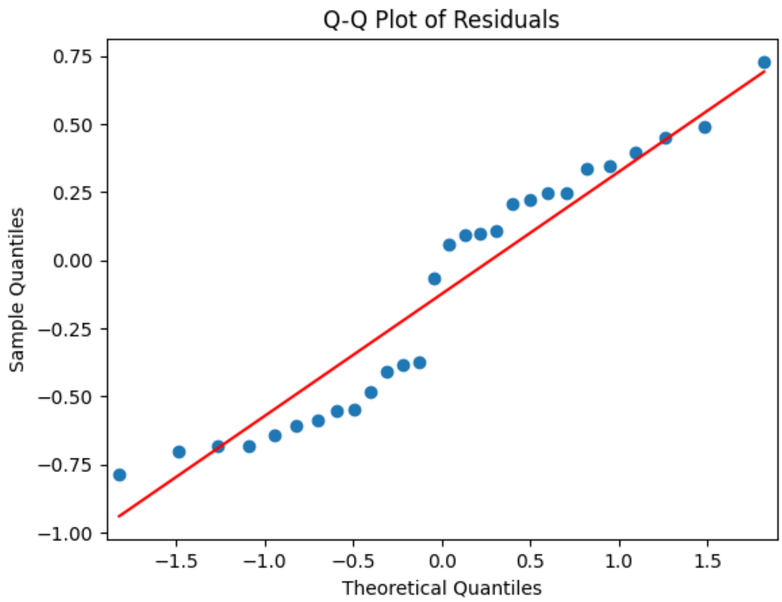


Fig. 4.10: Q-Q Plot of the ML algorithm.

The Figure 4.10 is a Q-Q plot which stands for Quantile-Quantile plot. This plot graphically represents the normality for the dataset. A dataset exhibits normality when the points are distributed in a way that resemble Gaussian normal distribution. The Q-Q plot is used as a tool to depict the normality. It provides visual comparison of how closely the data aligns with the distribution. The points almost replicate a linear line which implies the observed data follows the expected distribution.

4.0.4 Summary

This section showed the experimental results of the algorithms which were applied in the project. It consists of all the informative graphs and plots of the recommendation system model and mixed cropping ML- based model.

CHAPTER 5

Conclusion and Future Scope

5.1 Conclusion

We used a recommender system algorithm in order to recommend crops to the farmers based on the soil nutrient attributes such as N, P,K and pH. Soil sensors are deployed in the farm and the readings are transmitted via the cellular network to the cloud. The data from the cloud is procured and then inputted into the algorithm via the blockchain-based system which ensures the traversal of data securely without any intrusion. The algorithm outputs the recommended crops which can be grown on the farm. Furthermore, the crops acquired are used as an input in a mixed-cropping selection model, where it passes through ML algorithms and provides the most beneficial crop combinations. The choice of crops affects the yield and hence with the accuracy of our algorithm, such crop combinations are suggested which promise a healthy harvest and profit for farmers.

5.2 Future Scope

The objective of the project is to help farmers to optimize their yield and optimize the ways in which they can grow their crops. the future scope of the project can be considered in the following areas:
Employment of sensors and utilizing real time data helps the farmers in precision farming which in turn helps in increasing their yield and hence escalating their profit margin. An easy interface of an app helps farmers to understand and analyze the data provided by app easily.

5.3 Summary

This section concludes the project and stated all the problems that were considered and their respective solutions that were proposed. This chapter also talked about the future enhancements of the project.

REFERENCES

[1] A. Gupta, D. Nagda, P. Nikhare, and A. Sandbhor, "Smart crop prediction using iot and machine learning," *International Journal of Engineering Research & Technology (IJERT)*, pp. 2278–0181, 2021.

[2] B. P. Bhuyan, R. Tomar, T. P. Singh, and A. R. Cherif, "Crop type prediction: A statistical and machine learning approach," *Sustainability*, vol. 15, no. 1, 2023.

[3] B. Panigrahi, K. C. R. Kathala, and M. Sujatha, "A machine learning-based comparative approach to predict the crop yield using supervised learning with regression models," *Procedia Computer Science*, vol. 218, pp. 2684–2693, 2023. International Conference on Machine Learning and Data Engineering.

[4] U. Ahmed, J. C.-W. Lin, G. Srivastava, and Y. Djenouri, "A nutrient recommendation system for soil fertilization based on evolutionary computation," *Computers and Electronics in Agriculture*, vol. 189, p. 106407, 2021.

[5] M. Paul, S. K. Vishwakarma, and A. Verma, "Analysis of soil behaviour and prediction of crop yield using data mining approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 766–771, 2015.

[6] K. Bakthavatchalam, B. Karthik, V. Thiruvengadam, S. Muthal, D. Jose, K. Kotecha, and V. Varadarajan, "Iot framework for measurement and precision agriculture: Predicting the crop using machine learning algorithms," *Technologies*, vol. 10, no. 1, 2022.

[7] J. Lin, Z. Shen, A. Zhang, and Y. Chai, "Blockchain and iot based food traceability for smart agriculture," in *Proceedings of the 3rd international conference on crowd science and engineering*, pp. 1–6, 2018.

[8] M. Ur Rahman, F. Baiardi, and L. Ricci, "Blockchain smart contract for scalable data sharing in iot: A case study of smart agriculture," in *2020 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)*, pp. 1–7, 2020.

[9] R. Chaganti, V. Varadarajan, V. S. Gorantla, T. R. Gadekallu, and V. Ravi, "Blockchain-based cloud-enabled security monitoring using internet of things in smart agriculture," *Future Internet*, vol. 14, no. 9, 2022.

[10] I. Prieto, C. Violle, P. Barre, J. L. Durand, M. Ghesquière, and I. Litrico, "Complementary effects of species and genetic diversity on productivity and stability of sown grasslands," *Nature Plants*, vol. 1, no. 4, pp. 1–5, 2015.

[11] M. Li, J. Zhang, S. Liu, U. Ashraf, B. Zhao, and Q. Shuqing, "Mixed-cropping systems of different rice cultivars have grain yield and quality advantages over mono-cropping systems," *Journal of the Science of Food and Agriculture*, vol. 99, 12 2018.

[12] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," in *2016 Eighth International Conference on Advanced Computing (ICoAC)*, pp. 32–36, 2017.

[13] J. Nasar, A. Alam, A. Nasar, M. Z. Khan, *et al.*, "Intercropping induce changes in above and below ground plant compartments in mixed cropping system," *Biomed. J. Sci. Tech. Res*, vol. 17, pp. 13043–13050, 2019.

Report.pdf

ORIGINALITY REPORT

12%	9%	8%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	www2.mdpi.com Internet Source	1%
3	www.apaari.org Internet Source	1%
4	Bharati Panigrahi, Krishna Chaitanya Rao Kathala, M. Sujatha. "A Machine Learning-Based Comparative Approach to Predict the Crop Yield Using Supervised Learning With Regression Models", Procedia Computer Science, 2023 Publication	1%
5	www.coursehero.com Internet Source	1%
6	www.semanticscholar.org Internet Source	1%
7	Usman Ahmed, Jerry Chun-Wei Lin, Gautam Srivastava, Youcef Djenouri. "A nutrient recommendation system for soil fertilization	<1%