Migration Letters

Volume: 20, No: S13(2023), pp. 572-582

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)
www.migrationletters.com

Monitoring System for Increasing Crop Yield Production

G. Gifta Jerith[1], N. Noor Alleema[2], R. Umanesan[3] and S. Jagadeesh[4*]

Abstract

For humanity to survive, agriculture is by far the most crucial aspect. Low-yield crops are currently causing farmers problems. It might result in a food shortage. Low crop yield is mostly caused by a lack of understanding of soil fertility and crop choices. The same crop is chosen each season, which reduces the fertility of the land. The key to enhancing agricultural output is choosing the right crop, which can be done by doing a soil study and taking into account metrological parameters. This paper aims to generate smart farming and lower agricultural risk. This system depends on a monitoring system that uses IoT devices to gather data, and a machine learning algorithm to analyze the data once it has been extracted in real time. The data is collected that is pertinent to the plant using a number of sensors, and then we further analyze it using machine learning algorithms. Numerous methods, including KNN, Decision Tree, Random Forest, Gaussian Naive Bayes, and Extreme Gradient Boosting, are utilized for crop analysis. It can forecast the best crop to plant in order to increase yield production with the use of IoT and machine learning technology.

Keywords: *Internet of Things (IoT), Machine Learning (ML), Naive Bayes (NB), Decision Tree (DT), and Extreme Gradient Boosting (XGB) are some terms used in this context.*

Introduction

Agriculture, as we all know, is the main factor for food source, and for the vast majority of people in India, it is a only source for India's income. The Indian economy depends on it. Agriculture is essential to human survival. Given that it accounts for 17% of the GDB overall and employs more than 60% of the population, agriculture sector is a significant sector of the Indian economy. The previous few decades have seen significant expansion in Indian agriculture. The focus of autonomous technologies in the 21st century is on IOT (Internet of Things), and ML (Machine Learning) However, it is challenging to apply this technology in the agriculture sector. Increasing crop production is a current difficulty that requires a solution due to the constant changes in the physical and chemical parameters of the environment. Nowadays, the majority of agriculture is struggling with issues including crop failure and low yield due to a lack of nutrients and favorable environmental conditions. Agriculture is experiencing an increasing number of issues with extreme weather events that cause significant agricultural output losses as a result of the abrupt changes in the climate.

[1] Assistant Professor, Department of CSE (Artificial Intelligence and Machine Learning), School of Engineering, Malla Reddy University, Hyderabad, India. drg_giftajerith@mallareddyuniversity.ac.in

[2] Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India. noorashik21@gmail.com

[3] Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India. umanesan.r@gmail.com

[4*] Associate Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India. jagadeesh15.sj@gmail.com

IoT is crucial to smart agriculture in the present day. IoT sensors with the ability to provide data on agricultural lands. IOT technology aids in gathering data regarding environmental factors such as soil fertility, moisture content, and temperature. It might aid at the beginning of a plant's improved growth. The water level, humidity, temperature, and moisture level are all checked when the IOT-based crop monitoring system first turns on. It notifies the phone through an alert of the levels. The water pump will automatically sense and activate if the sensor detects a drop in water level. The LCD module shows all the data. This can also be observed in IOT, where data on humidity, moisture, and water level are displayed along with date and time based on the minute. Depending on the crops that are grown, the temperature can be controlled at a specific level. So, if we want to stop the water firmly, we may do so using the IOT, which has a button that allows us to forcefully stop the water.

Making algorithms that enable computers to independently learn from data and previous experience is a major focus of the branch of AI known as "Machine Learning." The development of machine learning has allowed computers to acquire new skills without human intervention. The main causes of agricultural production failure and low yield are inadequate nutrition levels and favorable environmental factors. Therefore, offering a solution for agricultural research is the main reason for this endeavor. For this, the data produced by the Internet of Things approach is put into the ML algorithms. Here, we use the IOT to produce data collection first, and the obtained data is then further analyzed by machine learning models. The information includes elements that include environmental physical characteristics including soil temperature, soil humidity, and light intensity. Our dataset's target variable is the rate at which height and width are increasing, which reflects the plant's growth rate. In order to give the plant, the best environment for optimum growth, the data is then exposed to the supervised machine learning algorithm.

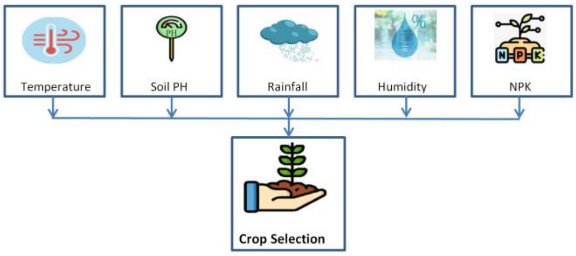

Figure 1: Crop selection factors

As shown in Figure 1, crop selection takes into account metrological parameters such rainfall, PH value, temperature, potassium (K), phosphorus (P), humidity, and nitrogen (N). Plant growth and productivity are directly impacted by meteorological elements [1–3].

Literature Review

Researchers Shafiullah Sharif et al. help farmers choose better crops by considering a variety of geographical and environmental parameters [4]. By aiding farmers in selecting the right crop and supplying information that farmers often lack, this technology lowers crop failure and productivity. Numerous techniques, including Decision Trees (DT), Random Forest Classifiers (RF), Neural Networks (NN), Gradient Boosting (GB), and Navies Bayes, are used in this. Gradient boosting is more accurate in this situation.

Amaan Choudhary et al. find the perfect conditions for the plant's growth, this method operates on sugarcane plants. [5] The major objective of the proposed system is to collect multiple sensor readings from the soil and predict the crop that would grow best in that specific type of soil. A variety of techniques are used in this, including artificial neural networks, feed-forward algorithms, backpropagation algorithms, decision trees, random forests, and linear support vector classifiers. By providing the climatic and physical circumstances, we can anticipate the growth of sugarcane plants by gathering all the data and training the model. We additionally identify the most ideal environmental factors for certain crops like millets, chilies, tea etc. We could create a dataset utilizing IoT that would inform us of the best plant to grow in a given soil and climate condition.

Richa Sing et al. find the ideal conditions for plant growth, research on a marigold plant was conducted [6]. The goal of this effort is to encourage smart farming and reduce risk in the agricultural industry. Numerous techniques, including the Gaussian Navies Bayes classifier, the linear support vector classifier, the decision tree, the random forest, the gradient boosting classifier, the logistic regression, and the stochastic gradient descent, are used in this. We've discovered the most ideal growing environment for marigold plants. In this, we employ gradient boosting classifier, linear SVC, and logistic regression, which are the best-fit methods with 83.33% accuracy.

Mayank Champaneri et al. uses a variety of machine learning approaches to primarily forecast the agricultural production [7]. Among the classifier models used here, random forest has the highest level of accuracy, followed by logistic regression and navies bayes. Farmers can choose which crop to produce to maximise production by using machine learning algorithms' forecasts, which consider many factors such as area, temperature, rainfall, and other features. This connects the agricultural and technological industries.

Amna Ikram et al. proposed a machine learning (ML) model, which is depends on information about soil and metrological parameters. [8] These variables include potassium (K), soil temperature, soil moisture content, phosphorus (P), CO2, pH value, nitrogen (N), Environmental Condition, and rainfall. Due to restricted resources, existing IoT-based solutions are less effective than our suggested model. Considering these variables, Real time data is transmitted to the cloud for analysis in the suggested model. The Android app also displays the findings. SCS combines the five ML algorithms listed below to improve efficiency and precision: Support Vector Machine, K-Nearest Neighbor, Random Forest Classifier (RF), Decision Tree, and Gaussian Naive Bayes. The Bahawalpur Agricultural Department provided a dataset with historical data for the previous fifteen years, which was used to predict rainfall. With the use of this dataset and the machine learning technique of Multiple Linear Regression, it is possible to predict future rainfall, which is crucial information for the health of any crop. The model has been trained to predict 11 crops, and its accuracy ranges from 97% to 98%.

Dhruvi Gosail et al. developed an IoT and ML system that is based on identifying and tracking the features of the soil and uses sensors to enable soil testing. [9] This method helps maintain crop health while reducing the chance of soil deterioration. The temperature, humidity, soil moisture, pH, and NPK nutrients of the soil are all monitored by this system using a variety of sensors. These sensors include ones for pH, soil temperature, and soil moisture. The microcontroller records the information gathered by these sensors and then analyses it using ML techniques like a Random Forest classifier to suggest the suitable crop to grow. In order to evaluate whether a plant may get a disease, this project's methodology predominantly uses a convolutional neural network.

Rohit Kumar Rajak et al. proposed the method is characterized by the use of a soil database gathered from farms, crops supplied by agricultural specialists, and soil testing lab datasets to reach metrics like soil. [10] The recommendation system will use the data from the soil testing lab to gather data, create an ensemble model using majority voting, and employ a various ML algorithm to recommend a crop efficiently and accurately for a relevant factor.

According to Prof. Meena Ugale et al., India's soil has been utilised for thousands of years, which has led to tiredness and nitrogen and mineral depletion, which reduces crop productivity [11]. In addition, the absence of contemporary applications creates a need for precision agriculture. The practice of precision agriculture, also called satellite farming, involves the use of various instruments and tactics for farm management that are centred on the observation, measurement, and response to crop variability both within and across fields. The suggestion of certain crops is a key use of precision agriculture. The result is higher agricultural yields and more revenue. Reviewing and assessing the efficacy of various methods for crop recommendation systems is the aim of this study.

Machine learning (ML) is a crucial viewpoint that Anakha Venugopal et al. suggested be used to provide a realistic and feasible solution to the crop yield problem [12]. Presently used methods mix manual counting with climate-smart pest control and satellite images, making the findings unreliable. For the purpose of agricultural output forecasting, this research employs a number of machine-learning techniques. The study's most accurate classification model was Random Forest, followed by Logistic Regression and Naive Bayes. Predictions made by machine learning algorithms will help farmers choose the crop to sow for optimal yield by considering factors like area, rainfall, temperature, and other features. Here, the IT and agricultural industries meet.

Choosing crops based on features from soil analyses and real-time sensing data is a significant advancement in smart agriculture research. [13] Three modules comprised a model suggested by Bhojwani et al.: crop maturity, crop management, and crop selection. For improved agricultural selection and health monitoring, they combined meteorological conditions with metrics for soil moisture, temperature, humidity, air pressure, and air quality. Real-time sensory data was analysed using the KNN algorithm on the ThingSpeak application.

Patil et al.[14] presented a scientific method for crop selection that combined real-time data collection with a microcontroller and a variety of temperature, soil, humidity, and infrared sensors. A number of data mining techniques are used to preprocess data and compare it with learnt data in order to forecast crops. The National Commodity and Derivatives Exchange's reported agricultural prices were also considered for crop forecasting. Applying the KNN classifier to analyse data.

Majumdar et al. have concentrated on Internet of Things-based agricultural techniques for weather tracking. [15] The prediction approaches are examined from both scientific and commercial standpoints, taking into account the cost of IoT components, security risks, and the impact of meteorological conditions on agricultural irrigation.

Imran proposed [16] an intelligent crop selection and irrigation system that takes into account many factors such as soil moisture content, temperature, humidity, and light intensity. Five different kinds of soil were used for the experiments: loamy, black, laterite, alluvial, and silt soil. from experiments indicate that crop selection may be based on the properties of the soil on various types of land. When analysing data, the ThinkSpeak application comes in handy. Also, farmers will be notified via an Android app of the ideal water level for their crops.

An Internet of Things architecture was presented by Rekha et al. [17] to enhance agricultural practices and maximise crop yield and profit by optimising land use. To collect data for various parameters and enable effective field monitoring, a wireless sensor network was set up in the field. The soil nutrients that aid in choosing the necessary fertiliser were discovered using the pH sensor. Additionally, an Android app was created to use meteorological data to assist farmers make informed irrigation choices.

Mulge et al. [18] introduced a crop prediction approach that uses machine learning (ML) algorithms to analyse real-time data of meteorological parameters, such as temperature, humidity, solar light, and precipitation, in order to maximise crop output and quality.

[19-21] Information leaks and security vulnerabilities are identified at the network layer of the Internet of Things. [22] A model for yield prediction by crop selection for sowing has been suggested by Paul et al. In order to analyse soil, an external dataset is obtained. Various soil micro and macronutrients are taken into account. To train the dataset, ML classification techniques (KNN and Naïve Bayes) are applied quickly and minimally. [23] Forecasting rainfall is another possibility. For validation, a metrological dataset with monthly rainfall statistics is utilised. [24] Four factors are taken into consideration in this work: rainfall, pH,

temperature, and humidity, in order to pick crops. It also provides guidance on how many nutrients are needed for a certain crop. Decision trees are used for dataset categorization, while SVM is used to forecast rainfall.

Objective

Reducing agricultural dangers, promoting smart farming methods, and increasing crop yield production are the objectives of the project effort. The main industry in India is agriculture. Many individuals in India find work in this industry, and it plays a significant role in the country's economy. These days, farmers often choose the wrong crop for a given soil type. Farmers in India are losing a lot of money because of this since it affects agricultural output. Nowadays, there are a number of factors that need to be considered while growing a certain crop on a particular soil type. An internet-of-things (IoT) monitoring system keeps tabs on how various physical elements, such temperature, humidity, soil nutrients, and moisture, affect plant development. Data received for the plants via various sensors and subsequently analysed using machine learning techniques. In order to improve agricultural yield output, this process may be used to build a system that will advise you on the best crop to grow.

Problem Statement

There are numerous issues affecting agriculture today, such as the impact of climate change and harsh weather that led affect the crops yield. Degraded soils are deficient in key nutrients, have poor structure, and are prone to erosion. These conditions have the potential to impede plant growth and diminish crop output. Climate change, such as temperature changes, changing rainfall patterns, and extreme weather events, can have a negative impact on agricultural productivity. Crop varieties may find it difficult to adjust to these changes, resulting in lower yields. Inadequate knowledge of contemporary farming techniques, insufficient training, and limited access to innovative agricultural technologies can all stymie productivity gains. Because of the low production and inability to repay their loans, farmers are killing themselves. Lack of nutrients in the soil and favorable environmental factors are the main reason of crop failure and low yield. The nutrients of the soil, the moisture, and environmental factors including metrological factors are the most significant variables that affect crop output. In earlier systems, the yield was predicted by analyzing the parameters of a single crop. Some systems use a small set of parameters, which results in inaccurate predictions. Repeated planting of the same crop reduces soil fertility.

Existing System

In the existing system, they select a single crop such as Sugarcane or marigold and take the parameters for these crops and analyze the environmental and metrological factors for the selected crops and then predict the result that how much yield it will give. It results this system will give the information about the selected corps and gives information about how can we increase the yield for the selected crop. This system would not fit for any other crops. In some systems, they have taken only limited parameters, information about the crops and analyze the factors and predict the result. But this kind of system will not give an accurate result. In some systems, they use many algorithms but that doesn't give good accuracy.

Proposed System

With this technique, a machine learning model gathers and analyses data from a real-time monitoring system to suggest the best crop for the given circumstances. An Internet of Things (IoT)-based monitoring system is used to track the effects of physical elements on plant growth, including temperature, humidity, soil nutrients, and moisture. The data was gathered for the plants using a variety of sensors, and machine learning techniques were employed to further analyse the data that was obtained. The data comprises rainfall, PH value, temperature, potassium (K), phosphorus (P), humidity, and nitrogen (N). And these data further undergo to a pre-processing technique called Exploratory Data Analysis. In EDA, data preprocessing includes categorical or numerical data analysis, data visualization, and statistical decision-making. And then important features are taken by feature selection method. This will use these data to develop a machine learning model by training the model that advises the optimum crop for maximizing yield.

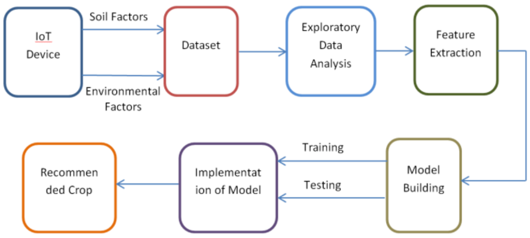

Figure 2: System Architecture

The India Meteorological Department gathers and reports rainfall data for Tamil Nadu from the CEIC. [25] The data that was gathered was unprocessed. The average rainfall for each month in each season for the previous four years is the basis for this statistics. Figures 3, 4, 5, and 6 depict the average rainfall in Tamil Nadu from 2020 to 2023, accordingly.

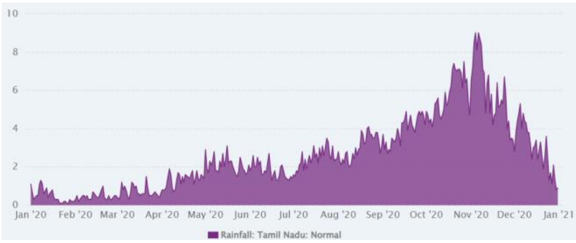

Figure 3: Average rainfall in the year 2020

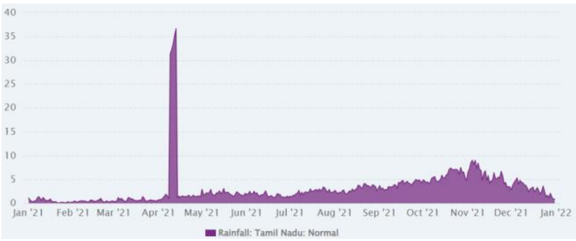

Figure 4: Average rainfall in the year 2021

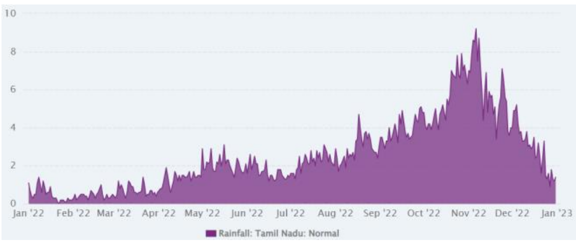

Figure 5: Average rainfall in the year 2022

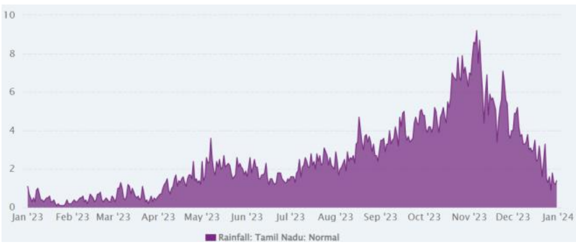

Figure 6: Average rainfall in the year 2023

In Tamil Nadu, the months of October, November, and December have more average rainfall than the other months. For crop maximization, soil fertility is also crucial, but it mostly depends on the availability of water resources.

Methodology

The following describes the comprehensive, step-by-step process used to create this suggestion system for boosting agricultural yield production:

Dataset Collection

The required dataset is a comma-separated values file format which including soil factor and environmental factors that has to be further pre-processed. The soil dataset taken into consideration for use in the specific proposed study is made up mostly of soil physical and chemical parameters as well as climatic information. The dataset was obtained from the Kaggle website and is publicly available. Eight key features and 2200 data points or examples are included in the dataset. 22 distinct crop types are included in this dataset: kidney beans, pigeon peas, moth beans, mung beans, black gramme, lentil, pomegranate, rice, maize, chickpea, banana, mango, grapes, watermelon, cotton, jute, coffee, and muskmelon.

Exploratory Data Analysis

One of the pre-processing techniques called as Exploratory Data Analysis (EDA). It is a technique for reviewing datasets in order to highlight their key properties. EDA frequently employs statistical graphics and other techniques for data visualization. EDA is different from conventional hypothesis testing in that it focuses mostly on what the data might reveal to us outside of formal modelling. It is possible to use or not use a statistical model. EDA differs from initial data analysis (IDA), which is more focused on completing gaps or missed values, modifying variables as necessary, and establishing the validity of the assumptions needed for model fitting and hypothesis testing. IDA is a part of EDA.

Feature Selection

Choosing the relevant features for the model is known as feature selection. A feature is a quality that has an impact on or contributes to the resolution of a problem. Each and every machine learning procedure depends on feature engineering, which basically entails the two processes of feature extraction and feature selection. Despite their shared objectives, feature selection and extraction methods are distinct. The key distinction between the two is that feature selection uses the dataset collection to determine important and relevant traits, while feature extraction adds new features. By narrowing down the data set to only the most relevant features, feature selection helps keep overfitting to a minimum by reducing the model's input variable.

Model Development

An ML model is built by collecting data from a variety of trustworthy sources, preparing it for modeling, choosing a modelling technique, developing the model, calculating performance measures, and choosing the best-performing model. Here in this system, it contains various machine learning algorithms were employed as supervised learning techniques.

K-Nearest Neighbor

The supervised learning technique known as K-Nearest Neighbour (KNN) is among the simplest machine learning algorithms. By comparing the newly collected data with the preexisting data, the K-NN algorithm determines which group the new data belongs to based on how similar it is to the current groups. Using the current data's similarity to classify a new data point, the K-NN approach preserves all previously classified data. It implies that the K-NN method may be able to rapidly classify new data into a suitable group at the time of arrival. This implies that current data may be properly and swiftly classified using the K-NN approach. Because the approach stores the training dataset instead of immediately learning from it, it is often referred to as a lazy learner. Alternatively, it uses the dataset to do a job while categorising data. In order to classify incoming data into a category that is highly linked to the existing data, the KNN approach just stores the dataset throughout the training phase.

This method computes distances using Euclidean geometry. The shortest path in any dimension between any two places is the Euclidean distance. The Euclidean distance is the most widely used method of measuring distance. The distance between two points on a plane with coordinates (A1, B1) and (A2, B2) may be calculated using the equation:

$$\text{Distance} = \sqrt{((A2-A1)^2 + (B2-B1)^2)} \quad (1)$$

Decision Tree

One kind of supervised learning is the Decision Tree algorithm. It is not parametric. Decision tree classifiers employ the greedy approach. It makes use of supervised learning, where the class labels and traits are used to represent a tree. To learn decision rules from training data and forecast the class or value of target variables, we must first train the system using a decision tree. Nodes and leaves are the two primary groups of the decision tree. Training the system using a decision tree is the initial step in learning decision rules from training data and forecasting the class or value of target variables. Each sub-tree with roots at the new nodes goes through this recursive procedure once more.

Entropy

Entropy is an expression of confusion or uncertainty, and the purpose of machine learning models and Data Scientists broadly is to decrease uncertainty. Entropy is determined using equation 2 as mentioned below:

$$F(T) = I_E\,(p1, p2, \ldots, p_J) = -\sum_{i=1}^{J}(p_i log_2 p_i) \quad (2)$$

Gini Impurity

Gini Impurity is a method that is used to design Decision Trees to identify how each feature of a dataset should be divided into nodes in order to create the tree. Gini impurity is calculated using equation 3 as shown below:

$$F(T) = I_M\,(p_{1,}p_{2,\ldots\ldots}p_J) = 1-\sum_{i=1}^{J}(p_i^2) \quad (3)$$

Random Forest Classifier

As a supervised learning technique, Random Forest may be used for both classification and regression. Decision trees are constructed in large numbers during the training period. The results are then separated into regression and classification outputs based on the number of classes. The forecast becomes more precise as there are more trees. The dataset includes variables including productivity, perception, temperature, and rainfall. The training procedure makes use of these dataset components. Only a portion of the dataset is taken into account. The remaining dataset forms the basis of the experiment. The random forest method is defined by three parameters: m, which specifies the amount of elements to take into account at a node split, and n tree, which defines the amount of trees that must be grown. We should make a certain amount of observations based on the node size, n terminal nodes.

The mechanism applies to decide how the decision tree nodes on the branch are ordered, is often utilized when constructing Random Forests based upon categorization data.

$$\text{Gini} = 1-\sum_{i=1}^{H}\,(p_i)^2 \quad (4)$$

Entropy, in contrast to the Gini index, requires more processing because it utilizes a logarithmic function to estimate the probability of a specific outcome to choose which way the node should branch.

$$\text{Entropy} = \sum_{i=1}^{H} - p_i * log_2(p_i) \quad (5)$$

Naïve Bayes Classifier

Order difficulties and multi-class arrangement problems are tackled with Naive Bayes. When given input values that may be represented as either yes or no, the naive Bayes method is very easy to implement. The Naive Bayes classifier recognises that the existence of one class member does not guarantee the presence of any other element. The NB classifier, which is grounded on Bayes theory, works well with data sources that have a high dimensionality. Spam filtering is only one of many uses for Naive Bayes. It can also provide real-time forecasts, estimate the likelihood of different classes of a target characteristic, and, when combined with collaborative filtering, improve recommendation system results. Class probability, which accounts for the likelihood of each attribute in the dataset, must be computed first. In order to determine the conditional probability of each information value, we look at the conditional probability assigned to each class value.

$$P\left({}^{B}/_{D}\right) = \frac{P\left({}^{D}/_{B}\right) * P(B)}{P(D)} \quad (6)$$

Extreme Gradient Boosting

EG Boosted decision trees are perform using EG Boosting. It's a kind of software library that was primarily created to increase speed and model performance. Recently, it has dominated the field of applied machine learning. Many Kaggle competitions have a pronounced dominance of Extreme Gradient Boosting models. This technique sequentially generates decision trees. In Extreme Gradient Boosting, weights are significant All independent variables are assigned weights, and this information is then utilized to inform the decision tree that predicts results. Prior to proceeding to the second decision tree, we assign more weight to factors that the previous one mispredicted. After that, a robust and accurate model is generated by merging all of these classifiers and predictors. Regression, classification, ranking, and personalised prediction are just a few of the problems that equation 7 may help with.

$$L^{(x)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(x-1)} + f_x(z_i)\right) + \Omega(f_x) \quad (7)$$

Results

Preprocessing is a common practice for the acquired data. Separating the training and test sets of data is the next step after pre-processing. Only 20% of the 2200 samples are used for testing purposes; the other 80% are used for training. Every sample is trained and tested using one of the following methods: K-Nearest Neighbour, Decision Tree, Random Forest, Naive Bayes, or Extreme Gradient Boosting. According to Table 1 and Figure 7, out of all the strategies tested, Extreme Gradient Boosting had the highest accuracy rate of 99.31%. Because of its adaptability, Python is used for implementation tasks. The projected harvest is shown in Figure 8.

Table 1: Algorithm vice Accuracy Result in Percentage

Algorithm	Accuracy
K-Nearest Negighbor (KNN)	97 %
Decision Tree (DT)	90 %
Random Forest Classifier (RF)	99 %
Naïve Bayes Classifier (NB)	99 %
Extreme Gradient Boosting (XGBoost)	99.31 %

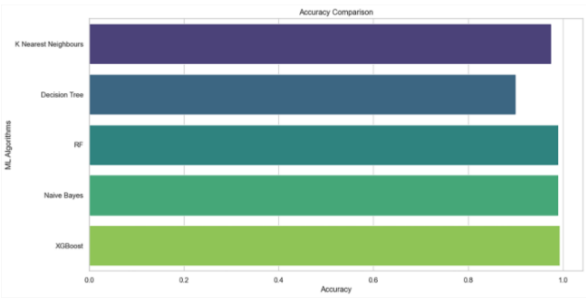

Figure 7: Accuracy Comparison between Algorithms

```
In [52]: save_model(model, 'model.pkl')

In [53]: prediction = model.predict((np.array([[80,70,80,300,80,7,250]])))
         print("The suggested crop for given climatic Condition is: ",prediction)

         The suggested crop for given climatic Condition is:  ['papaya']
```

Figure 8: Recommended Crop

Conclusion

Farmers that use conventional farming practices have issues like low crop production brought on by unpredictable weather, the use of incorrect fertilizers and water, and inappropriate crop choices. The previous study only employed a few criteria, which are insufficient for high crop production. Our research aims to choose the best crop in order to increase crop output. Compared to earlier studies, the machine learning (ML) algorithms used in this proposed study offer greater result at a lower computational cost. The calculation of accuracy uses a variety of machine-learning techniques. Crop recommendation is done by the Random Forest classifier. Implemented a technique to anticipate crops using data gathered in the past. The proposed model enables farmers to decide what crop to grow on the field. Farmers will be able to increase agricultural output, decrease chemical use, make better use of water resources, increase crop yields, and halt soil erosion on agricultural land with the aid of our endeavour.

Future Work

Additional functionality may be added to the system by enhancing it:

1. The primary goal of the next effort is to create an enhanced dataset with more characteristics.
2. Our goal should be to create a model that can detect whether crop leaves are healthy or sick and, if a disease is present, which disease it is.
3. Create user-friendly mobile apps and websites.

References

H. A. Issad, R. Aoudjit, and J. J. Rodrigues, "A comprehensive review of data mining techniques in smart agriculture," Engineering in Agriculture, Environment and Food, vol. 12, no. 4, pp. 511–525, 2019.

J. Flak, "Technologies for sustainable biomass supply-overview of market offering," Agronomy, vol. 10, no. 6, 2020.

P. M. Kavita, "Satellite-based crop yield prediction using machine learning algorithm," in 2021 Asian Conference on Innovation in Technology (ASIANCON), pp. 1466–1470, PUNE, India, August 2021.

Shafiullah Sharif, Pushpa H, Shwetha RB, Ramya O G, Pooja K R.: Crop recommendation using machine learning technique. International Journal of Engineering Research & Technology (IJERT), Volume 10, Issue 1, (2022).

Amaan Choudhary, Azharsaied Qureshi, Bhaskar Ananthraman, Prof. Sonali Deshpande.: AI and IoT Based Monitoring System for Increasing the Yield in Crop Production Security. International Research Journal of Engineering and Technology (IRJET), Volume: 07, Issue: 12, (2021).

Richa Singh, Sarthak Srivastava, Rajan Mishra.: AI and IoT Based Monitoring System for Increasing the Yield in Crop Production. International Conference on Electrical and Electronics Engineering (2021).

Mayank Champaneri, Darpan Chachpara, K.J. Somaiya, Mansing Rathod, Chaitanya Chandvidkar.: Crop Yield Prediction Using Machine Learning. International Journal of Science and Research (IJSR) (2021).

Amna Ikram, Waqar Aslam, Roza Hikmat Hama Aziz, Fazal Noor, Ghulam Ali Mallah, Sunnia Ikram, Muhammad Saeed Ahmad, Ako Muhammad Abdullah, and Insaf Ullah.: Crop Yield Maximization Using an IoT-Based Smart Decision. Hindawi Journal of Sensors, 2022 (2022).

Dhruvi Gosai1, Chintal Raval, Rikin Nayak, Hardik Jayswal, Axat Patel.: Crop Recommendation System Using Machine Learning. International Journal of Scientific Research in Computer Science, Engineering and information Technology, Volume 7, Issue 3, [p.554-557] (2021).

Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, Avinash Devare.: Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique" - International Research Journal of Engineering and Technology (IRJET), 04, Iss 12, (2020).

Prof. Meena Ugale, Nimeesha Venkatavelu, Pranay Patil, Suraj Rane.: Review of Machine Learning Techniques for Crop Recommendation System. International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 10, Issue III, (2022).
Anakha Venugopal, Aparna S, Jinsu Mani, Rima Mathew, Prof. Vinu Williams.: Crop Yield Prediction using Machine Learning Algorithms. International Journal of Engineering Research Technology (IJERT), Volume 9, Issue 1, (2021).
Y. Bhojwani, R. Singh, R. Reddy, and B. Perumal, "Crop selectionand IoT based monitoring system for precision agriculture," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 2, 2017.
J. S. Raj, International Conference on Mobile Computing and Sustainable Informatics, EAI/Springer Innovations in Communication and Computing, Springer Cham, 2020.
P. Majumdar, S. Mitra, and D. Bhattacharya, "IoT for promoting agriculture 4.0: a review from the perspective of weather monitoring, yield prediction, security of WSN protocols, and hardware cost analysis," Journal of Biosystems Engineering, vol. 46, no. 4, pp. 440–461, 2021.
S. Imran, "Effective crop selection and conservative irrigation using IoT," International Journal of Science and Research (IJSR), 2016.
P. Rekha, V. P. Rangan, M. V. Ramesh, and K. V. Nibi, "High yield groundnut agronomy: an IoT based precision farming framework," in 2017 IEEE Global Humanitarian Technology Conference (GHTC), vol. 2017, pp. 1–5, San Jose, CA, USA, Dec. 2017.
M. Mulge, M. Sharnappa, A. Sultanpure, D. Sajjan, and M. Kamani, "An invitation to subscribe: the international journal of analytical and experimental modal analysis," vol. 10, no. 1, pp. 1112–1117, 2020.
W. U. Khan, X. Li, A. Ihsan, M. A. Khan, V. G. Menon, and M. Ahmed, "2021. NOMA-enabled optimization framework for next-generation small-cell IoV networks under imperfect SIC decoding," IEEE Transactions on Intelligent Transportation Systems, pp. 1–10, 2021.
W. U. Khan, Z. Ali, M. Waqas, and G. A. S. Sidhu, "Efficient power allocation with individual QoS guarantees in future small-cell networks," AEU-International Journal of Electronics and Communications, vol. 105, pp. 36–41, 2019.
W. U. Khan, J. Liu, F. Jameel, M. T. R. Khan, S. H. Ahmed, and R. Jantti, "Secure backscatter communications in multi-cell NOMA Networks: enabling link security for massive IoT networks," in In IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), vol. 2020, pp. 213–218, July 2020.
M. Paul, S. K. Vishwakarma, and A. Verma, "Analysis of soil behaviour and prediction of crop yield using data mining approach," in Proceedings -2015 International Conference on Computational Intelligence and Communication Networks, CICN, vol. 2016, 766–771, Jabalpur, India, 2015.
N. Gnanasankaran and E. Ramaraj, "A multiple linear regression model to predict rainfall using Indian meteorological data," International Journal of Advanced Science and Technology, vol. 29, no. 8, pp. 746–758, 2020.
M. Mulge, M. Sharnappa, A. Sultanpure, D. Sajjan, and M. Kamani, "Agricultural crop recommendation system using IoT and M.L," The International Journal of Analytical and Experimental Modal Analysis, vol. 12, no. 6, 2020.
https://www.ceicdata.com/en/india/rainfall-by-states/rainfall-tamil-nadu-normal
Kaggle, https://www.kaggle.com
Wikipedia, https://www.wikipedia.com
GitHub, https://www.gitgub.com