

Smart Farming Using Machine Learning and Deep Learning

P. Darwin, M. Jaya Raju,
K. Naveen Babu, Y. Lavanya, Y. Satya Sai
Electronics and Communication Engineering
Godavari Institute of Engineering & Technology (A)

Abstract

Using the new technologies that have emerged in agriculture at the start of the fourth industrial revolution is known as "smart farming." Crop Recommendation, Fertilizer Recommendation, and Crop Leaf Disease Detection are the three essential components that make up the platform. Farmers enter critical information about the soil, such as location and N-P-K levels. The system uses climatic data from a Weather API and machine learning algorithms to process this data. Machine learning techniques anticipate and suggest the best crops to grow in a particular area based on environmental parameters and nutrient levels. Users indicate the crop they intend to grow and supply information about the nutrients in their soil. Advanced machine learning algorithms examine the soil's nutrient content and compare it to the needs of a certain crop. The technology addresses nutrient surpluses or shortfalls by producing customized fertilizer recommendations. This guarantees that fertilizer is used optimally, meeting the specific requirements of the soil and the selected crops. This feature sets machine learning apart by offering a complete crop health and productivity solution. By using a deep learning model—more precisely, the ResNet CNN algorithm—for crop leaf disease identification, it surpasses traditional techniques. Farmers can input photos of sick crop leaves, and the system uses deep learning's potent skills to precisely identify the disease. In addition, the platform provides recommendations and treatments for the illnesses identified, giving farmers practical knowledge to protect their harvests.

Keywords

Machine learning, ResNet CNN, Crop Recommendation, Fertilizer Recommendation, Leaf disease detection

I. INTRODUCTION

One of the most common vocations in our nation is agriculture. A large number of individuals rely on agriculture for their livelihood. Every farmer faces significant challenges at every level of the agricultural industry. Modern technologies and numerous methods are employed these days to streamline agricultural labour. Intelligent farming relies heavily on machine learning, which assists farmers in making informed judgments and, inadvertently or directly, contributes to better agricultural decision-making. Incorrect crop selection can result in lower yields and financial losses for farmers. Occasionally, crops are damaged by unforeseen cyclones. This leads to farmers giving up on agriculture and moving more frequently from rural to urban areas in search of a better life. Some farmers also commit

suicide as a result of crop damage, but technology is beginning to address these issues. The suggested system is put into practice using machine learning and deep learning techniques. It provides crop recommendations based on values for nitrogen, phosphorus, and potassium. It also uses state and city data to determine temperature conditions and recommends the best crop based on that data and machine learning algorithms. and by entering the N, P, and K values as well as the crop the farmer want to grow in his field, it suggests fertilizer with a low composition in that specific soil to nurture the particular crop. It can make recommendations for how to prepare the land for growing a specific crop by using machine learning algorithms to analyze the crop and its N, P, and K values. Crop disease identification is another crucial phase in agriculture. Farmers cannot identify illnesses by visual inspection alone. Therefore, by just submitting photos of the sick plant leaves, we can detect sickness using deep learning techniques, particularly the ResNet CNN algorithm. We may anticipate high productivity if all these procedures are followed correctly, and machine learning in conjunction with agriculture can streamline the industry's conventional procedures. Modern machine learning and deep learning techniques have simplified the agricultural process. The farmer's productivity increases with proper crop recommendations, and the amount of fertilizer they apply is reduced based on the composition of the soil. Farmers also spend less on agriculture thanks to these clever techniques, which also reduce unnecessary fertilizer expenditures. Crop to crop has a different ratio of N, P, and K values, and the right environmental conditions are also necessary for plants to grow properly. Some crops grow only in high humidity circumstances like wheat and some plants need extremely low temperatures like apple and some crops need more water to grow for example rice and it is very necessary to think about the temperature conditions of that region to produce the finest crop. Another crucial component of agriculture is production.

II. LITERATURE SURVEY

Palaniraj, A. et al (2021, Apr) proposed a machine learning-based crop and fertilizer recommendation system prediction model. This model examines a number of variables, including soil pH value and location. In addition to those variables, they recommend crops and fertilizers based on the composition of phosphorus, potassium, and nitrogen. The Support Vector Machine approach, which splits the mixed data into distinct classes and plots each data item as a point in n-dimensional space, was utilized along with preprocessed crop data to provide an accuracy of 90.01%.

Devdatta A. Bondre, et al. (2019, Sep) proposed a system of crop yield and fertilizer recommendation using machine learning algorithms. This paper focuses on creation of prediction model which may be used for prediction of future crop yield. They used different machine learning techniques evaluated in agriculture for estimating future year crop prediction based on the previous years data. This is achieved by using machine learning algorithms like support vector machine and Random forest on agricultural data and recommends fertilizer suitable for particular crop. The accuracy of Support Vector Machine algorithm is 99.47% for yield prediction and Random forest accuracy is 97.48%.

Prof. Kiran Sowmwanshi, et al proposed a system of crop prediction and fertilizer recommendation using machine learning. They also used the factors like magnesium, Calcium, Lime, carbon etc. By investigating the composition and other traits of the soil, it is envisioned that the right crop and fertilizer would be forecast as part of their proposed work. They used the attributes like Location, weather, Area, Productivity for predicting crop yield. They also used Support Vector Machine algorithm for predicting crop yield.

Archana Chougale et al suggested creation of a crop compatibility and fertilizer recommendation system based on ontology. Additionally, this system makes fertilizer recommendations based on the soil's NPK (nitrogen, phosphorus, and potassium) composition and makes use of ontology-stored research data from previous years. It offered recommendations for fertilizer as well as advice on which crops would be best in a given area. Along with Ontology, it makes use of the Random Forest method for crop recommendations and the K Means Clustering technique for fertilizer recommendations.

Naga V Vara Prasad Mella et al proposed a system of crop yield prediction and fertilizer recommendation using Voting based Ensemble classifier. They use the attributes like location, pH value from which nature of the soil is determined and it uses the amount of nutrients like Nitrogen(N), Phosphorus(P), Potassium(K). Along with these it uses the location, weather API and amount of rainfall This system analyses and tarin the data with Support Vector Machine, Random Forest, ,K Nearest Neighbour and Voting classifier for creating model. This system predicts crop yield and suggests fertilizer ratio. It also recommends pesticides based on types of pest.

Taranjeet Singh et al proposed a system Crop and fertilizer recommendation and plant disease diagnosis system using machine learning and Internet of Things. In this system they develop an IOT device that analyses the quality of soil and it also predict crop diseases by scanning the leaves of crop. It also measures pH of the soil and it estimates the Nitrogen, Phosphorus and Potassium content of the soil based on PH value. This system uses XG Boost algorithm mainly for classification purpose to predict crops which has 99% accuracy rate, Based on N,P,K values it recommends fertilisers.

Suresh Singh Rajpurohit et al presented a machine learning-based method for Growfarm Crop, Fertilizer, and Disease Prediction. This system uses soil and pH data as inputs to generate a website that predicts which crops will grow best in the soil and which fertilisers can be used to treat plant diseases. It predicts crop, fertilizer, and leaf diseases using a variety of machine learning algorithms, including Decision Tree, Random Forest, Logistic Regression, Gaussian Naïve Bayes, Support

Vector Machine, and ResNet CNN.

Dr.B. Ratnakanth et al suggested a deep learning-based method for predicting diseases and recommending crops and fertilizers. The suggested methodology in this project helps farmers maximize crop productivity and recommends the most lucrative crop for a given area. In order to maximize agricultural output and so help satisfy the nation's growing demand for food supplies, it offers crop selection based on economic and environmental conditions. It also indicates the ideal time to apply fertilizers and makes crop recommendations based on factors like rainfall, temperature, season, and kind of soil.

Virendra Panpatil et al. It had created a framework for generating profitable yield proposals, which had done a significant amount of work for Indian ranchers. They created a framework including classifier models such the Naive Bayes Classifier, KNN, and Decision Tree Classifier. For the purpose of determining the best time to sow, grow, and harvest plants, the suggested framework can be utilized. In order to achieve greater exactness, they employed a unique classifier; for example, Nave Bayes outperforms decision trees in datasets with more variability in terms of exactness. The framework's greatest advantage is that, after everything is said and done, it can be incredibly useful and versatile.

III. EXISTING SYSTEM

Up until now, research has only focused on crop prediction using various soil properties and machine learning techniques. It also recommends fertilisers based on crop predictions and predicts crop yield using data on soil nutrients and crop yield and using some Machine learning algorithms. Finally, it recommends fertilisers for selected crops based on a variety of datasets, including location, crop yield, and fertiliser data and in existing models it predicts fertilisers for the recommended crop. In leaf disease prediction the existing systems predicts diseases of the leaf and it did not explain about the causes of the disease

IV. PROPOSED SYSTEM

In the proposed system we are extended the crop recommendation deeply to predict crop by using information soil nutrient values like Nitrogen, Phosphorus ,Potassium, Rainfall, pH value, State and City details and it connects to the weather API to know about the climatic conditions in that area and we trained the machine by using six different algorithms(Decision tree, Random forest, Naïve Bayes, Logistic Regression, XG boost and Support vector machine and choose the best one to predict the crop. Among these XG Boost gives the more accuracy. In addition to crop prediction the proposed system predicts the amount of calories and amount of nutrients present in that recommended crop. And it also giving information about the recent years production on that crop for unit area and it is our first module

In fertiliser recommendation we ask the farmers about which crop they want to cultivate the system takes his soil nutrient data like Nitrogen, Phosphorus and potassium values and it take the name of the crop the farmer wants to cultivate. It uses big data. Leveraging some machine learning algorithms and it cross references the soil nutrient data with the specific

crop requirements and it predicts the fertiliser and provided suggestions on how to prepare the land to cultivate that particular crop.

Our third module is the crop leaf disease prediction using deep learning model In order to detect agricultural leaf disease, our project uses deep learning models, more precisely the ResNet CNN algorithm. This distinguishes it from conventional systems that might rely on less complex techniques. Crop disease leaves can be uploaded by farmers, and the system can identify the illness and offer practical treatment recommendations in real time, allowing for prompt and efficient intervention. Our proposed system can also identifies the causes of the disease.

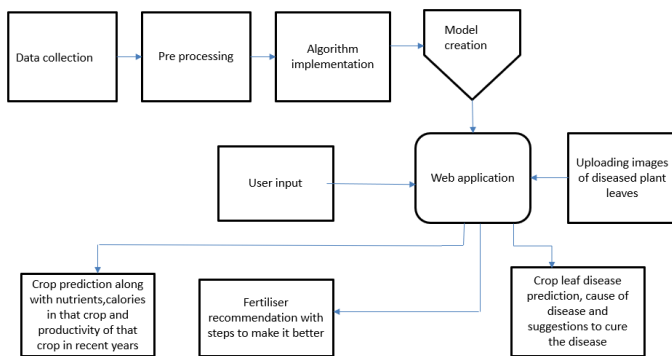


Fig. 1 : Block diagram of Proposed model

V IMPLEMENTATION

A) Crop Methodology:

The CSV file "crop-recommendation.csv" contains the dataset that is imported into the model for training and testing. It has 2200 instances and 8 features, such as temperature, humidity, rainfall, pH level, nitrogen (N), phosphorus (P), and potassium (K), as well as the goal variable "label," which represents the crop. To comprehend the structure of the dataset, exploratory data analysis, or EDA, was conducted. Analysis is done on descriptive statistics, data types, and unique values for the target variable. Heatmaps and other visualization tools are used to examine the relationship between features.

The dataset is divided into the target variable (Y) and features (X). Features are chosen based on how well they suit crop growth. One-hot encoding, also known as label encoding, is used to translate category labels into a numeric representation.

A number of machine learning models, including Decision Tree, Naive Bayes, Support Vector Machine (SVM), Logistic Regression, Random Forest, and XGBoost, are selected for cross-validation. Every model is trained on the training set and evaluated on the test set.

The dataset is divided into training and testing sets using the train_test_split technique set and assessed on the test set. Every model is trained using the available data. With Decision Tree Classifier(criterion="entropy", random_state=2, max_depth=5), for example, a decision tree that is constrained. Using the testing set, each model's accuracy, precision, recall, and F1-score are assessed. Cross-validation scores are calculated to guarantee the models' robustness.-With the pickle module, trained models are

stored for later use. Every model has a.pkl file. A barplot is designed to evaluate the differences in the models' accuracy. Among the models, XG Boost exhibits the highest accuracy. The Random Forest model is used to produce predictions using sample data points. For example, a set of soil and environmental characteristics is used to predict the most appropriate crop.

B) Fertilizer Methodology:

It starts by reading two CSV files into Pandas DataFrames (cropandfert), cpdata.csv and Fertilizer.csv. Both Data Frames' column names are standardized by being converted to lowercase and having spaces removed. The string replacement is used to standardize certain crop names in the FertDataFrame. The code then uses the unique crop names retrieved from the FertDataFrame to filter data from the crop DataFrame. A function change_case is defined to reduce the case of the captions in both Data Frames. The standardized and filtered data frames are saved as MergeFileCrop.csv and file MergeFileFert.csv.

Users can input crop name, nitrogen, phosphorous, and potassium levels by using the Flask web app Nutrient (N), Phosphorous (P), and Potassium (K) values for the designated crop are retrieved by the user from the fertilizer dataset (fertilizer.csv). N, P, and K are used to calculate the difference between user input levels and dataset levels. A key is chosen from a predetermined dictionary (fertilizer_dic) to provide a recommendation based on the identification of the biggest difference. The fertilizer dictionary includes suggestions for various situations involving the amounts of N, P, and K in the soil. Suggestions include types of fertilizers to use, soil amendments to consider, and certain activities to follow depending on the amounts of nutrients. The user is shown the recommendation that is shown in web interface and it provides suggestions to prepare the land better.

C) Methodology for Disease Prediction:

Collect a varied collection of plant photos that include both healthy and unhealthy plants. Make sure the dataset includes a range of crops and illnesses. Divide the dataset into sets for testing, validation, and training.

Clean up the data to address any irregularities in the dataset. To ensure model compatibility, resize photos to a uniform size. Add more data if necessary to make the training set more diverse. The process of Exploratory Data Analysis (EDA) involves visualizing the dataset's class distribution. Examine a few sample pictures to learn what makes healthy and unhealthy plants different.

Select a deep learning architecture (like ResNet) that is appropriate for picture classification. Train and deploy the model with PyTorch or any deep learning framework of your choice. Put training pipelines and data loaders into place. Assess the trained model using the validation dataset. Track performance indicators including F1 score, recall, accuracy, and precision. Adjust the model as needed in light of validation results. Test various hyperparameters to enhance model efficiency. Employ strategies like weight decay, gradient cutting, and learning rate scheduling.

Assess the generalization performance of the final model by

evaluating it on the test set. Look for indications of either an underfit or an overfit. Construct a Flask web application to facilitate user communication. Provide an easy-to-use interface for uploading photos of plants. Use the trained model to process uploaded photos by implementing the backend logic. Gather user input regarding anticipated outcomes. Take into account putting in place a feedback loop to make the model better over time.

Select an appropriate hosting platform to install the Flask application. Install the program and make sure it can be accessed online. Track usage of the program and gather input. Regularly add fresh data to the model to enhance performance.

D) Algorithms used:

Decision Tree Classifier:

Among the most effective supervised learning techniques for classification and regression applications is the decision tree. It constructs a tree structure that resembles a flowchart, with each internal node signifying a test on an attribute, each branch designating a test result, and each leaf node (terminal node) containing a class name. When a stopping criterion—such as the maximum depth of the tree or the minimum number of samples needed to split a node—is satisfied, the training data is recursively split into subsets based on the values of the attributes. Depending on a metric like entropy or Gini impurity, the Decision Tree algorithm chooses the appropriate attribute to partition the data depending on during training.

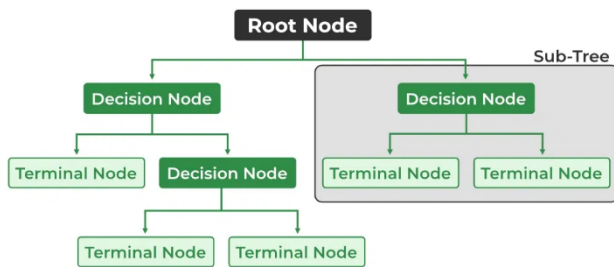


Fig 2: Decision Tree Algorithm

Naive Bayes :

A group of classification algorithms based on Bayes' Theorem are known as naive Bayes classifiers. It is actually a family of algorithms rather than a single method, and they are all based on the same principle—that is, each pair of features being classified stands alone. Let's start by thinking about a dataset. The Naïve Bayes classifier, one of the most straightforward and efficient classification algorithms, facilitates the quick creation of machine learning models with quick prediction capabilities.

Random forest:

The random forest approach consists of an ensemble of decision trees, where each tree is composed of a bootstrap sample, or sample of data taken from a training set with replacement. One-third of the training sample is designated as test data; this is referred to as the out-of-bag (oob) sample, and it is something we will discuss more. Feature bagging is then used to introduce yet another randomization, increasing dataset variety and decreasing decision tree correlation. The prediction's

determination will change depending on the kind of problem. The individual decision trees in a regression job will be averaged, and in a classification work, the predicted class will be determined by a majority vote, or the most common categorical variable.

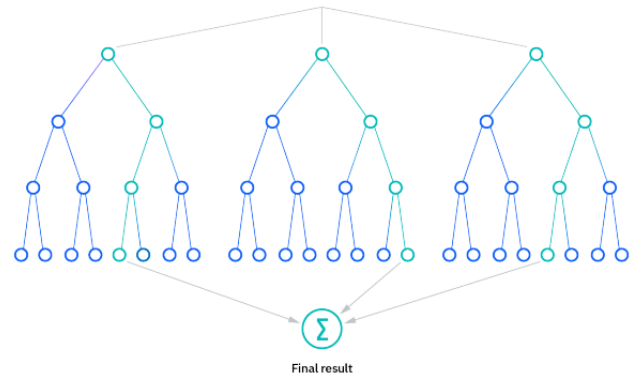


Fig 3 Random Forest Algorithm

Logistic Regression:

For categorization and predictive analytics, this kind of statistical model—also referred to as the logit model—is frequently employed. Logistic regression uses a dataset of independent variables to estimate the likelihood of an event occurring, such as voting or not. Because the result is a probability, the dependent variable has a range of 0 to 1. A logit transformation is performed to the odds in logistic regression, which are the probability of success divided by the probability of failure. This logistic function is also referred to as the log odds or the natural logarithm of odds. It is expressed by the following formulas.

XG Boost:

XG Boost, which stands for "Extreme Gradient Boosting," is an optimized distributed gradient boosting library created for the effective and scalable training of machine learning models. It is an ensemble learning technique that combines the predictions of several weak models to generate a stronger prediction. Because it can handle large datasets and produces state-of-the-art performance in many machine learning tasks, including regression and classification, XG Boost has grown to become one of the most well-known and extensively used machine learning algorithms.

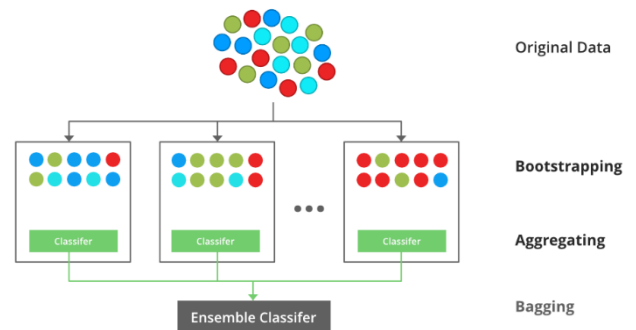


Fig 4: XG Boost Algorithm

Support vector machine:

A supervised machine learning approach called Support Vector Machine (SVM) is used for regression as well as classification. Even yet, classification problems are the most appropriate use

for regression problems. The SVM algorithm's primary goal is to locate the best hyperplane in an N-dimensional space that may be used to divide data points into various feature space classes. The hyperplane attempts to maintain the largest possible buffer between the nearest points of various classes. The number of features determines the hyperplane's dimension. The hyperplane is especially a line if there are just two input features. The hyperplane transforms into a 2-D plane if there are three input features. If there are more than three features, it gets hard to imagine.

ResNet CNN:

Residual Network: This architecture presented the idea of Residual Blocks as a solution to the vanishing/exploding gradient issue. We employ a method in this network called skip connections. By omitting some layers in between, the skip connection links activations of one layer to those of subsequent layers. Thus, a residual block is created. These leftover blocks are stacked to create resnets. The idea behind this network is to let the network fit the residual mapping rather of having layers learn the underlying mapping. So, let the network fit, rather than using, example, $H(x)$, initial mapping.

VI. RESULTS

Fig 5 web interface for crop recommendation

Better go with *groundnut* in your farm

Crop production for 2020-2021 is 3515 in india
 Crop production for 2021-2022 is 3699 in india
 Crop production for 2022-2023 is 3873 in india
 calories is Calories: 102 KCAL
 nutrients is [{"nutrientId": 1093, 'nutrientName': 'Sodium, Na', 'nutrientNumber': '307', 'unitName': 'MG', 'derivationCode': 'LCCS', 'derivationDescription': 'Calculated from value per serving size measure', 'derivationId': 70, 'value': 201, 'foodNutrientSourceId': 9, 'foodNutrientSourceCode': '12', 'foodNutrientSourceDescription': 'Manufacturer's analytical; partial documentation', 'rank': 5800, 'indentLevel': 1, 'foodNutrientId': 29725495, 'percentDailyValue': 25}, {'nutrientId': 1008, 'nutrientName': 'Energy', 'nutrientNumber': '208', 'unitName': 'KCAL', 'derivationCode': 'LCCS', 'derivationDescription': 'Calculated from value per serving size measure', 'derivationId': 70, 'value': 102, 'foodNutrientSourceId': 9, 'foodNutrientSourceCode': '12', 'foodNutrientSourceDescription': 'Manufacturer's analytical; partial documentation', 'rank': 300, 'indentLevel': 1, 'foodNutrientId': 29725489}, {'nutrientId': 1092, 'nutrientName': 'Potassium, K', 'nutrientNumber': '306', 'unitName': 'MG', 'derivationCode': 'LCCS', 'derivationDescription': 'Calculated from value per serving size measure', 'derivationId': 70, 'value': 66.0, 'foodNutrientSourceId': 9, 'foodNutrientSourceCode': '12'}

Fig 6: Crop prediction with amount of calories, Nutrients present in the recommended crop and production of recommended crop in recent years

Fig 7: Web interface for fertilizer recommendation

The K value of your soil is high.
 Please consider the following suggestions:

1. Ensure the soil is deeply aerated with a shovel, and thoroughly watered to aid in the dissolution of water-soluble potassium. Allow the soil to completely dry before repeating the process of digging and watering two to three times.
2. Rake through the soil to eliminate any rocks, using a soil sifter to remove as many as possible. Rocks containing minerals like mica and feldspar gradually release potassium into the soil over time through natural weathering processes.
3. Discontinue the application of potassium-rich commercial fertilizers. Instead, opt for fertilizers with a '0' in the final number field, indicating a lack of potassium. Alternatively, shift away from commercial fertilizers altogether and transition to solely using organic materials to enhance soil fertility.
4. Incorporate crushed eggshells, seashells, wood ash, or soft rock phosphate into the soil to introduce calcium. Additionally, blend in organic compost, comprising up to 10 percent of the soil mixture, to amend and harmonize soil composition.
5. Select NPK fertilizers with reduced potassium levels, and prioritize organic fertilizers as they typically have lower NPK values.
6. Cultivate a cover crop of legumes to facilitate nitrogen fixation in the soil. This approach satisfies the soil's nitrogen requirements without elevating phosphorus or potassium levels.

Fig 8: Fertiliser recommendation of farmer wanted crop with suggestions to prepare the land better

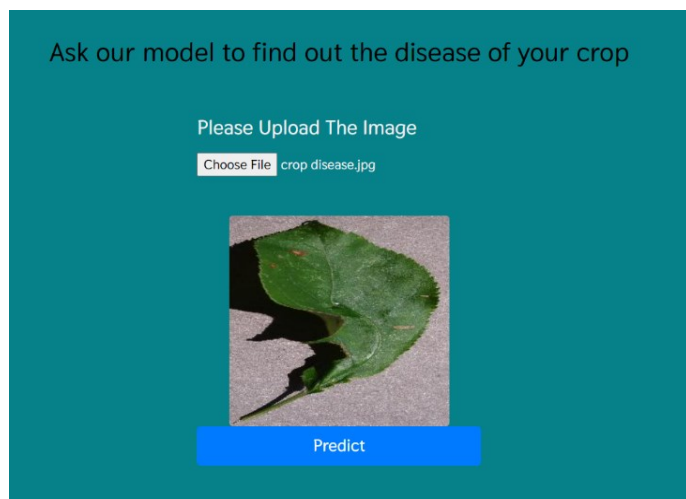


Fig 9: Web interface to upload the images of sickness plant leaf

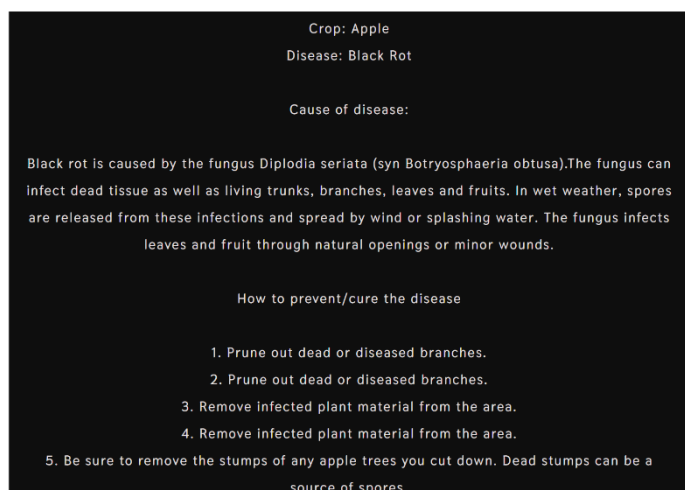


Fig 10: Disease prediction, causes of the disease and steps to cure the disease

VII CONCLUSION AND FUTURE WORK

The system developed using machine learning can predict the suitable crop for that field according to the soil nutrient composition and using rainfall, pH value of the soil and location details and it can also give information about the calories and top five nutrients present in that particular crop and also it can display recent years crop production in India. This information is very useful for every farmer especially it is very much useful for farmers who are starting the farming recently. It provides fertilizer recommendation by input the farmers interested crop and soil nutrients values in his land and it can predict the crop leaf disease and additionally it can also give information about the cause of the disease and it provides suggestions to cure that particular disease. To expand this work in future by using some IOT devices and combining it with the machine learning and it may give suggestions about the crop from the stage of cultivation to stage of harvesting and it may provide alerts at every challenge in agriculture and also it provides information about the Natural calamities like cyclones etc in advance and provide suggestions to protect the crop and it may reduce Human involvement also.

VIII REFERENCES

- 1) Palaniraj A, Balamurugan A S, Durga prasad R, Pradeep P (April 2021). Crop and fertilizer Recommendation System using Machine Learning. In 2021 International Research Journal of Engineering and Technology (IRJET) (pg.319-323).
- 2) Devdatta A. Bondre, Mr. Santosh Mahagoankar (September 2019). Prediction of crop yield and Fertilizer Recommendation using Machine Learning Algorithms. In 2019 International Journal of Engineering Applied Sciences and Technology (IJEAST) (Pg.371-376).
- 3) Prof. Kiran Somwanshi, Priyanka Rajendra Sonawane, Pooja Nagraj patil, Tushar Sanjay Lohar, Mayuri Sandeep Jadhav. Crop and Fertiliser Recommendation Using Machine Learning (March 2023). In 2023 International Journal of Engineering Research and Applications (IJERA) (pg.28-32).
- 4) Archana Chougale, Vijay Kumar Jha, Debajyothi Mukhopadhyay (July 2019). Crop Suitability and Fertilizers Recommendation using Data Mining Techniques. In 2019 https://www.researchgate.net (pg.978-981).
- 5) Naga V Vara Prasad Mella, Venkata Murali Pentakoti, Crop yield prediction and Fertilizer Recommendation using Voting Based Ensemble Classifier (August 2022). In 2022 Journal of Engineering Sciences. (pg.262-270).
- 6) Taranjeet Singh, Saurabh Anand, Anmol Sehgal, Siddhesh Mahajan, Prof. Pranoti Kavimandan, CROFED- Crop and Fertilizer Recommendation and Disease diagnosis system using Machine Learning and Internet of Things. In 2022 International Journal of Innovative Research in Technology (IJIRT) (pg 660-667).
- 7) Suresh Singh Rajpurohit, Sowmen Parul, Rakshitha Bhole (March 2023). GrowFarm-Crop, Fertilizer and Disease Prediction using Machine Learning. In 2023 International Research Journal of Engineering and Technology (IRJET) (pg 26-31).
- 8) Dr. B. Ratnakanth, N. Raj Kumar, Shruthi Prasad, Rampalli Dinesh Kumar, Sunil Varma, Nagulanived (March 2021). Crop and Fertilizers Recommendation and Disease Prediction using Deep Learning. In 2021 Dogo Rangsang Research Journal (pg 846-854).
- 9) Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate, "Crop Prediction System using Machine Learning Algorithms", International Research Journal of Engineering and Technology, Feb 2020.