International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 04 | Apr 2023 www.irjet.net p-ISSN: 2395-0072

SOIL FERTILITY AND PLANT DISEASE ANALYZER

Shivamkumar Prasad[1], Sandip Yadav[2], Nilesh Gahlot[3], Aadarsh Mishra[4], Sonali Padalkar[5]

[1]Student VIII SEM, B.E., IT Engineering, SLRTCE, Mumbai, India
[2]Student VIII SEM, B.E., IT Engineering, SLRTCE, Mumbai, India
[3]Student VIII SEM, B.E., IT Engineering, SLRTCE, Mumbai, India
[4]Student VIII SEM, B.E., IT Engineering, SLRTCE, Mumbai, India
[5]Assistant Professor, Department of Information Technology, SLRTCE, Maharashtra, India

---

Abstract - *Agriculture forms an integral part of our lives and is a major source of employment in India, with more than half of the population relying on it. It serves as the backbone of our economy, but the yield of crops depends on several factors, with soil quality and plant diseases being the most significant. Early detection of diseases is critical for achieving an efficient crop yield, as bacterial spots, late blight, Septoria leaf spots, and yellow-curved leaf diseases can all hurt crop quality. For better crop growth, it is imperative to have efficient soil fertility prediction and early plant disease analyzer systems in place. Additionally, automatic methods for classifying plant diseases help take prompt action upon detecting symptoms of leaf diseases. Improving crop yield prediction techniques can aid farmers and other stakeholders in making better decisions regarding agronomy and crop selection, taking into account factors such as temperature, humidity, pH, rainfall, and crop name from previous historical data. This system can provide an accurate status of plant diseases and recommend better-suited crops for the soil.*

***Key Words*: Agriculture, for farmers, Machine Learning, Crop recommendation, Real-time detection.**

1. INTRODUCTION

The proposed system would leverage advanced technologies such as machine learning artificial intelligence and data analytics to analyze large volumes of data and provide accurate recommendations to farmers the system would utilize historical data and real-time data from sensors installed in the fields to provide timely and accurate recommendations to farmers the system would be user-friendly and accessible through mobile and web applications making it easy for farmers to access the recommendations from anywhere at any time. By providing accurate recommendations to farmers the proposed system would not only help farmers make informed decisions about which crop to grow but also improve their yields and profitability the system would also contribute to the overall food security of the country by ensuring that farmers grow the most suitable crops and reducing the risk of crop failure due to plant diseases or soil fertility issues.

Overall, the proposed system has the potential to revolutionize the way farmers make decisions about which crop to grow and help them overcome the challenges associated with plant diseases and soil fertility issues with the right support and investment this system could be a game-changer for the Indian agricultural sector and contribute to the economic growth and development of the country

2. LITERATURE REVIEW

Paper1: Kiran Moraye, Aruna Pavate, Suyog Nikam and Smit Thakkar [7]

The research paper has utilized a 10-fold cross-validation technique to develop a model that can accurately predict the correlation between climate and crop yield. The accuracy of the model was found to be 87%, which is a promising result. However, the model only considered climate factors and did not account for other essential factors like soil quality, pests, and chemicals used, which significantly impact crop yield. Therefore, it is crucial to incorporate these factors into the model to develop a comprehensive decision-making tool that farmers can use to select the appropriate crop for their fields. The web application that the researchers propose will be an excellent tool for farmers and users to make better decisions based on the climate of a particular season. By providing recommendations for the most suitable crops based on the prevailing weather conditions, farmers can optimize their yield and reduce the risk of crop failure.

Furthermore, the application will also be useful for policy planners in areas like import-export, pricing, and marketing. By providing early predictions of the yield for different crops, the application can help policymakers make informed decisions even before the crop is harvested. In conclusion, while the research paper has provided a promising model for predicting the correlation between climate and crop yield, it is essential to incorporate other essential factors like soil quality and pests to develop a comprehensive decision-making tool. The proposed web application has the potential to be an invaluable tool for farmers and policymakers alike, providing early insights into the crop yield for different crops based on prevailing weather conditions.

Paper 2: Divyansh Tiwari, Mrityunjay Ashish, Nitish Gangwar [9]

The authors of this paper have utilized the concept of transfer learning to develop an automated system that can diagnose and classify diseases in potato leaves the system can detect early blight late blight and healthy leaves with high accuracy achieving a classification accuracy of 978 over the test dataset compared to similar studies this novel solution demonstrated a significant improvement of 58 and 28 over two previous works 2 and 3 respectively this remarkable performance suggests that the proposed technique can be highly effective in detecting diseases in potato leaves in their early stages enabling farmers to take preventive measures and enhance their crop yields by providing an accurate and efficient tool for diagnosing potato leaf diseases this technology can help farmers avoid significant crop losses due to disease early detection and intervention can prevent the spread of diseases and minimize the use of harmful chemicals which can be detrimental to both the environment and human health in conclusion the proposed system utilizing transfer learning is a promising approach for automating disease diagnosis and classification in potato leaves its high accuracy and efficiency can enable farmers to detect diseases early and take the necessary preventive measures ultimately enhancing their crop yields and improving overall agricultural productivity

Paper3: Shriya Sahu, Meenu Chawla, and Nilay Khare [6]

This study a comprehensive approach is taken to predict the most suitable crop based on various parameters ranging from soil characteristics to atmospheric conditions the authors consider a wide range of soil parameters such as type ph-level iron copper manganese sulfur organic carbon potassium phosphate and nitrogen in order to classify the dataset to classify the dataset the authors use the random forest algorithm which has been shown to provide good accuracy with a low error rate furthermore the proposed framework is designed to handle large datasets using the MapReduce programming model which can significantly improve the processing speed and efficiency of the system the different phases of the proposed work include data collection data classification using the random forest algorithm implementation on the Hadoop framework utilizing the MapReduce programming model and final prediction the implementation is carried out on ubuntu 1404 it's with Hadoop 260 and the dataset is collected from various online sources by utilizing this approach farmers can make more informed decisions about which crop to plant based on the specific characteristics of their soil and the prevailing atmospheric conditions this can help to maximize crop yield and ensure sustainable agriculture practices moreover the proposed framework has the potential to be applied to other regions and crops making it a versatile tool for enhancing agricultural productivity

3. SYSTEM ANALYSIS

3.1 Problem Definition

The agricultural sector in India is vital to the country's economy and employs a significant portion of the population the quality of the crops produced by farmers is dependent on various factors including soil quality and the prevention and identification of plant diseases that can cause significant damage to crops hindering their growth and yield and can have a significant impact on the farmer's livelihood detecting diseases early is critical to preventing their spread and an automated detection system can provide an efficient solution typically plant diseases are visible in various parts of the plant such as the leaves however manual diagnosis using photographs can be a time-consuming process therefore automated computational methods must be developed to detect and classify diseases based on leaf symptoms with an accurate and efficient disease detection and classification system in place farmers can provide the appropriate treatment to protect their crops and maximize their yields

3.2 Proposed System

Our proposed crop recommendation system addresses the common issues that farmers face in decision-making by utilizing various parameters like climate soil quality and location to provide accurate crop recommendations this system also incorporates an early detection mechanism for plant diseases allowing farmers to take prompt action before the disease causes significant crop damage additionally the system suggests the appropriate fertilizers for specific crops which will ensure optimal crop growth and yield by utilizing this system farmers can increase their crop production and profitability while also reducing the risk of crop loss due to factors like disease and inappropriate fertilization overall our proposed system is a reliable and effective solution to the challenges faced by farmers in crop planning and decision-making

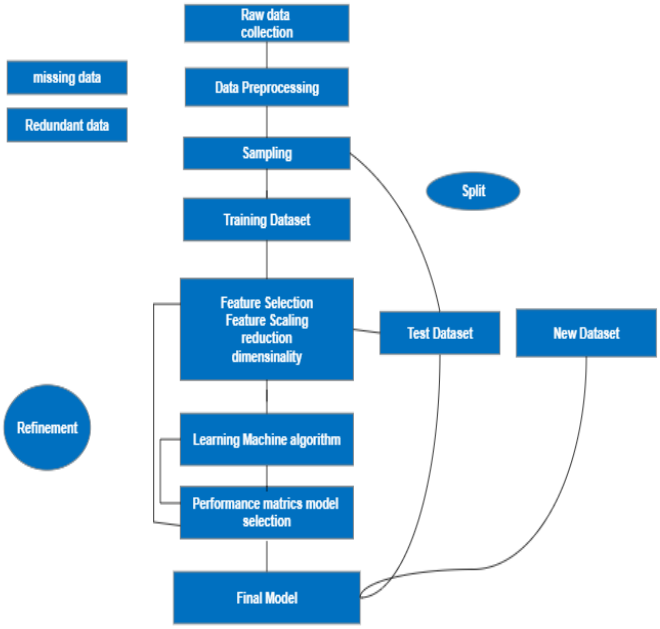

Fig.1 Flow Chart of the System

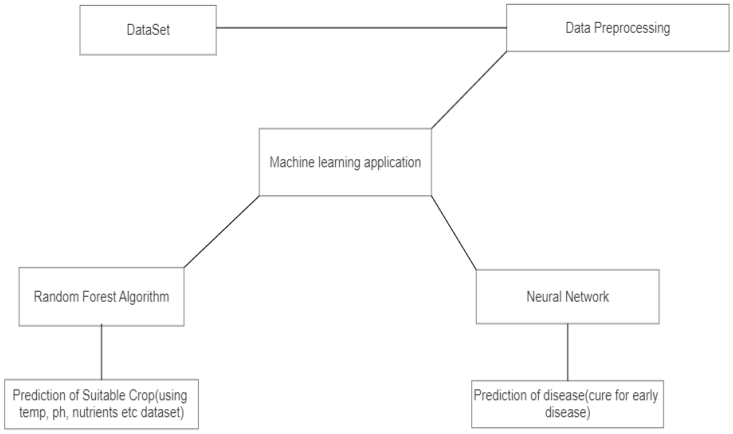

Fig.2 System architecture of the proposed system

3.3 Crop recommendation and plant diseases analyzer System Working:

Plant Disease Detection Process

The detection of plant diseases involves a systematic process consisting of four phases the first phase is the acquisition of images which can be done through digital cameras mobile phones or from the web in this phase images of plant leaves are gathered with desired resolution and size the second phase involves image segmentation where the image is simplified to become more meaningful and easier to analyze there are different segmentation methods such as k-means clustering convolution neural networks etc that can be applied the third phase is feature extraction where features such as color shape and texture are extracted from the segmented image to determine the meaning of a sample image after feature extraction the last phase is classification where the input image is classified as either healthy or diseased different classifiers such as k-nearest neighbor KNN support vector machines SVM artificial neural network ANN convolution neural network CNN nave bayes and decision tree classifiers have been used in the past CNN is the most commonly used classifier as it is simple-to-use and robust the success of the detection system largely depends on the efficiency of these four phases.

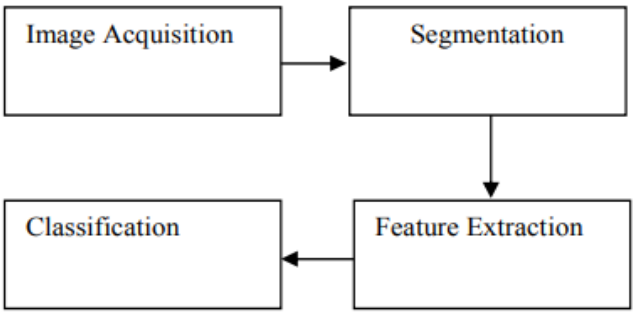

Fig.3 Phases of plant disease detection system

Image Acquisition

the collected images are processed to simplify their representation and make them easier to analyze this is achieved through feature extraction which involves various methods such as k-means clustering and convolutional neural networks to segment the images once the segmentation is done the features of the area of interest such as color shape and texture descriptors are extracted these features are then fed into a classifier which determines if the leaf is healthy or diseased and identifies the type of disease if any the classification is performed using different machine learning algorithms such as convolutional neural networks and support vector machines finally the system provides a diagnosis and recommended actions to prevent the spread of disease.

Image Segmentation & Feature Extraction

To extract features from the segmented image, the system analyzes various aspects of the image, such as the color histogram, texture, and shape descriptors. These features provide valuable information for classification and help distinguish healthy leaves from diseased ones. Once the features are extracted, they are fed into a classifier, which determines whether the leaf is healthy or diseased and if diseased, what type of disease it may have. This classification is done using various machine learning algorithms, such as convolutional neural networks and support vector machines. Finally, the system provides the user with a diagnosis and recommended course of action to prevent the spread of the disease.

Classification

The classification phase is a crucial step in the plant leaf detection system as it determines whether the input image is healthy or diseased if the image is found to be diseased the system may further classify it into various diseases based on the features extracted in the previous phase in the past researchers have utilized a variety of classifiers such as k-nearest neighbor KNN support vector machines SVM artificial neural networks ANN convolutional neural networks CNN naive bayes and decision tree classifiers each classifier has its own strengths and limitations which makes it more suitable for specific types of data and applications among these classifiers CNN is the most popular due to its simplicity and robustness it can automatically learn features from the input data and has shown remarkable performance in various computer vision tasks however researchers are continuously exploring new classifiers and hybrid models to improve the accuracy and efficiency of the plant leaf detection system by doing so they can provide a more effective and reliable solution to identify diseased plant leaves which can ultimately lead to better crop yield and improved food security

Crop Prediction and Suggesting Process

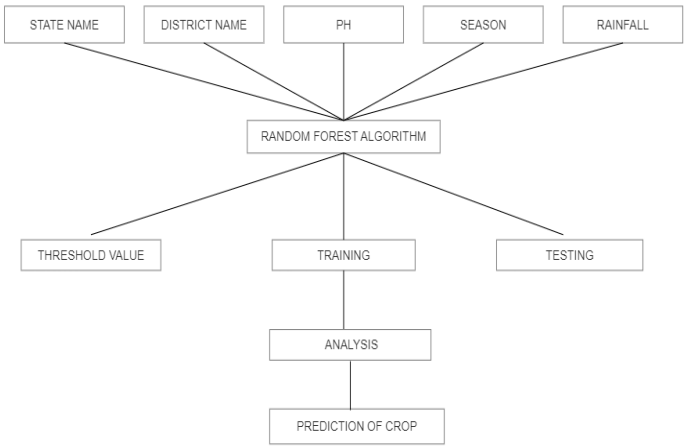

Fig.4 Crop prediction system

Our data management methodology involves a comprehensive process that involves collecting datasets from various sources and storing them in a designated data storage location to ensure the data is accurate and reliable we implement data cleaning data reduction and data normalization techniques to refine the datasets the data cleaning process eliminates inaccurate and incomplete data while data reduction simplifies numerical and alphabetical digital information. once we have a refined dataset we apply data normalization techniques to ensure that the numeric values are uniformly scaled across the entire dataset this process helps to prevent any inconsistencies in the data and ensures that it is easier to analyze we then use feature selection and data extraction techniques to identify the most relevant features for our project output and extract the necessary data for further processing this approach helps us to streamline the data processing and analysis phase as we can focus on the most pertinent information finally to enhance the accuracy of our predictions we utilize the random forest classifier a machine learning algorithm that is well-suited for predictive modeling with this approach we can generate the production output of the crop with a high degree of accuracy which is crucial for decision-making processes in agriculture.

A. Random Forest Classifier

The random forest algorithm also called the random decision forest classifier, is a popular technique for supervised learning tasks such as classification, regression, and association. This approach uses multiple decision trees during both training and testing, and the final classification or predictive regression is determined by the mode of the decision trees. This method helps prevent overfitting to the training and testing samples. As a type of supervised learning, the random forest algorithm maps input data to an output. It creates a forest with numerous decision trees, which results in more accurate and robust results. This feature makes it ideal for handling large datasets with high dimensionality, as it can handle complex and diverse data with ease. Another advantage of the random forest classifier is its ability to provide feature importance measures. By analyzing the decision trees in the forest, we can identify the input features that are most important in predicting the output. Overall, the random forest classifier is a powerful tool in machine learning, frequently used in various performance measures. Our system utilizes this algorithm to achieve accurate and reliable results in our predictions. It is a critical component of our data processing and analysis pipeline, playing a vital role in our success.

4. RESULTS AND DISCUSSION

Our project involves a preprocessing technique that includes a feature selection process to choose the most suitable features these selected features are then passed to the random forest classifier to classify and predict if the crop is suitable for the agricultural field and provide the expected crop production as output the accuracy of the output is evaluated to ensure the reliability of the results during the classification phase the input image is checked for any signs of disease and if any is found the image is further classified into specific diseases researchers have employed several classifiers

Fig 5. Accuracy Level Graph for Existing System and Proposed System

over the years such as k-nearest neighbor KNN support vector machines SVM artificial neural network ANN convolutional neural network CNN nave bayes and decision tree classifiers among these classifiers CNN has been the most commonly used due to its simplicity and robustness each classifier has its own set of advantages and disadvantages but CNN has proven to be a versatile and easy-to-use technique overall our project utilizes advanced techniques to provide farmers with accurate and reliable information about their crops which can significantly improve their productivity and profitability

5. CONCLUSION

The agricultural sector is an essential component of the economy and farmers face numerous challenges including crop selection which affects their productivity and profitability machine learning and data analysis are advanced technologies that can provide valuable insights to farmers and help them make informed decisions to prevent crop failure the proposed system employs convolutional neural network and random forest algorithm models to assist farmers in selecting the appropriate crop based on soil and atmospheric conditions with an accuracy rate of 8988 and 8826 respectively by expanding the dataset to include various crops and seasons the accuracy of the models can be further improved the systems web-based availability makes it easily accessible to millions of farmers providing them with valuable insights and recommendations to enhance their productivity and profitability furthermore people who want to set up a kitchen garden can also benefit from this system by receiving useful tips and advice the proposed system is an excellent example of how advanced technologies can be leveraged to improve society's welfare it has the potential to significantly enhance the agricultural industry's productivity and profitability thereby contributing to the country's economic development in conclusion the proposed system provides farmers with valuable insights and recommendations which can significantly improve their productivity and profitability the agricultural sector is critical to any country's economic development and the proposed systems potential impact cannot be overemphasized

REFERENCES

[1] Keerthan Kumar T G, Shubha C, Sushma S A," Random Forest Algorithm for Soil Fertility Prediction and Grading Using Machine Learning", IJITEE,2019.

[2] Dr. V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, AP. Janani," An Effective Crop Prediction Using Random Forest Algorithm", IEEE,2021.

[3] Pranay Malik, Sushmita Sengupta, Jitendra Singh Jadon," Comparative Analysis of Soil Properties to Predict Fertility and Crop Yield using Machine Learning Algorithms", 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence 2021), IEEE.

[4] Melike Sardogan, Adem Tuncer, Yunus Ozen," Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm",3rd International Conference on Computer Science and Engineering, IEEE,2020.

[5] Dharmesh Vadalia, Minal Vaity, Krutika Tawate, Dynaneshwar Kapse," Real Time soil fertility analyzer and crop prediction", International Research Journal of Engineering and Technology (IRJET),2017.

[6] Shriya Sahu, Meenu Chawla, Nilay Khare," An Efficient Analysis of Crop Yield Prediction Using Hadoop Framework Based on Random Forest Approach", International Conference on Computing, Communication and Automation (ICCCA2017).

[7] Kiran Moraye, Aruna Pavate, Suyog Nikam, Smit Thakkar," Crop Yield Prediction Using Random Forest Algorithm for Major Cities in Maharashtra State", International Journal of Innovative Research in Computer Science & Technology (IJIRCST),2021.

[8] Monali Paul, Santosh K. Vishwakarma, Ashok Verma," Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach", 2015 International Conference on Computational Intelligence and Communication Networks, IEEE.

[9] Divyansh Tiwari, Mrityunjay Ashish, Nitish Gangwar," Potato Leaf Diseases Detection Using Deep Learning", Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020), IEEE.