WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Agriculture Intelligence and Support System for Farmers

Prof. Shinde B. A[1], Pranav Kshirsagar[2], Piyusha Pol[2], Bhavika Tambe[2], Pratiksha Khaire[2]

[1]Professor, Department of Computer Engineering
[2]Students, Department of Computer Engineering
Shree Ramchandra College of Engineering, Pune, Maharashtra, India

ABSTRACT

The agricultural industry plays a crucial role in India's economic development and Gross Domestic Product. To promote agricultural modernization, agricultural informatization has become essential. With the advancement of agricultural infrastructure, utilizing modern information technology to deliver personalized agricultural information and timely recommendations has proven to be an effective solution. This article provides a comprehensive review of recent research on the application of intelligent recommender systems in agricultural information. It begins with an introduction to the content analysis method used to sort the papers. Next, it presents the background concepts of recommender systems and key technologies. The article then describes in detail the applications of recommender systems/technologies for agricultural information. Finally, it concludes with a summary and an outlook on the application of recommender systems for agricultural information. One major challenge faced by the industry is the lack of scientific approaches to soil fertility. Due to limited knowledge of soil nutrients, farmers often rely on myths and assumptions when starting their cultivations. This project aims to address this issue by suggesting the best crops based on soil fertility and providing a fertilizer plan to minimize the required amount of fertilizers. To achieve this, a cross-platform web application has been developed. It suggests the best crops according to available soil fertility and provides a fertilizer plan based on the Nitrogen (N), Phosphorus (P), and Potassium (K) values. This optimized fertilizer usage will increase profitability and prevent soil degradation.

Keywords: Crop Recommendation, Fertilizer Recommendation, Machine Learning, Agriculture.

INTRODUCTION

Agriculture plays a vital role in our country, contributing approximately 18% of the Gross Domestic Product (GDP) according to the annual report for 2018-2019 from the Department of Agriculture, Co-operation and Farmer Welfare. It is the most important sector of the Indian economy, employing nearly half of the country's workforce. Despite being the second-largest producer of fruits and vegetables globally, as reported by the Department of Agriculture, Co-operation and Farmer Welfare in 2018-2019, farmers continue to face financial distress due to crop losses. Inappropriate crop selection and fluctuating market prices due to surplus production are key factors contributing to this situation. To address these challenges, a system is needed to predict the best crop based on soil characteristics, environmental factors (such as temperature, humidity, and water level), and market demand. This project aims to provide such a system using technological advancements such as data analysis, cloud storage, and the Internet of Things.

METHODOLOGY

Interpreting Inputs for the Projects: The input design serves as a crucial link between the information system and the user. It involves developing specifications and procedures for data preparation, ensuring that transaction data is in a usable form for processing. This can be achieved through various methods, such as reading data from written or printed documents or directly inputting data into the system. The primary focus of input design is to control the amount of input required, minimize errors, avoid delays and extra steps, and keep the process simple. Additionally, input design aims to provide security, ease of use, and privacy.

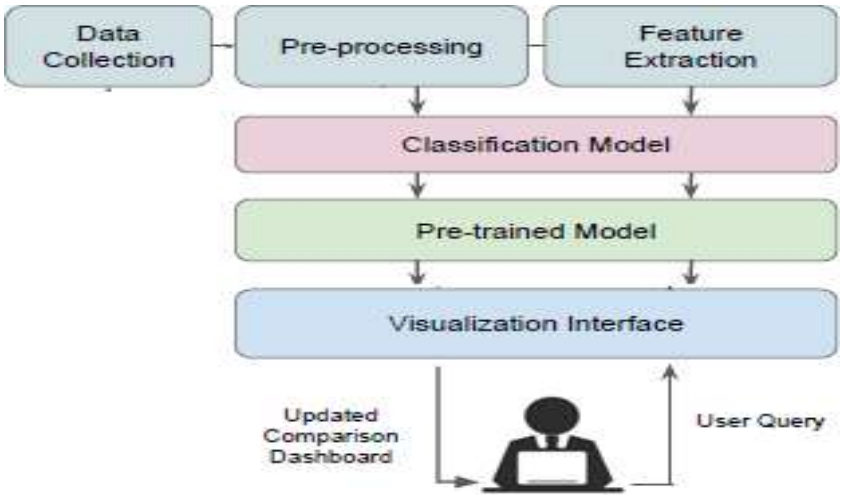
The input design process considers the following factors:

- What data should be provided as input?
- How should the data be arranged or coded?

- The dialog to guide the operating personnel in providing input.
- Methods for input validation and steps to follow when errors occur.

1. **Input Design:** Input design is the process of converting a user-oriented input description into a computer-based system. Its main objective is to avoid errors in the data input process and provide accurate information to the management through the computerized system. This is achieved by creating user-friendly screens for data entry, especially when handling large volumes of data. The goal is to make data entry easier and error-free, while also providing facilities for record viewing.
2. **Data Validation:** During data entry, the input design ensures the validity of the entered data. User-friendly screens are designed to facilitate data manipulation, and appropriate messages are displayed to guide the user when necessary. The objective is to create an input layout that is easy to follow and minimizes the chances of mistakes.

SYSTEM ARCHITECTURE



DATASET

The purpose of this dataset is to contribute to the development and deployment of automated machine learning algorithms for the classification of plant diseases. The ultimate goal is to achieve fast and accurate disease detection. Our plan is to continuously expand this dataset by adding more images, creating a larger and more comprehensive expert-annotated dataset for future Kaggle competitions. This will also allow us to explore advanced methods for disease classification and quantification.

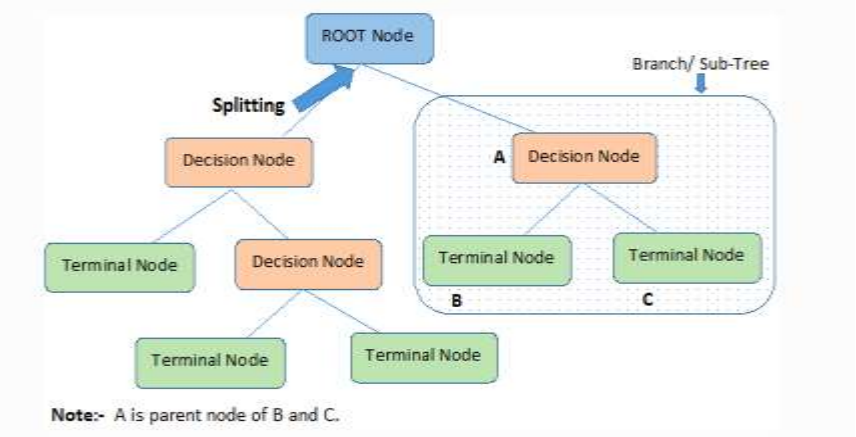
We have manually captured 3651 high-quality images of real-life symptoms of multiple apple foliar diseases. These images exhibit variations in illumination, angles, surfaces, and noise. A subset of the images has been expert-annotated to create a pilot dataset specifically for apple scab, cedar apple rust, and healthy leaves. This pilot dataset has been made available to the Kaggle community.

ALGORITHMS

DECISION TREE

The Decision Tree algorithm is a supervised learning algorithm used for classification and regression problems. It aims to create a training model that predicts the class or value of the target variable by learning simple decision rules from existing data. A Decision Tree is a tree-structured classifier, where internal nodes represent features of the dataset, branches represent decision rules, and leaf nodes represent the outcome. It follows a graphical representation to determine possible solutions based on given conditions. The tree is built using the CART algorithm (Classification and Regression Tree), which iteratively splits the tree based on the best features. Each split results in a decision node with multiple branches, leading to leaf nodes that contain the final outcomes."

Accuracy is 90.0 %



GAUSSIAN NAÏVE BAYES

Gaussian Naive Bayes is a variant of the Naive Bayes algorithm designed for handling continuous data that follows a Gaussian (normal) distribution.

Naive Bayes algorithms are a collection of supervised machine learning classification methods that rely on the principles of Bayes' theorem.

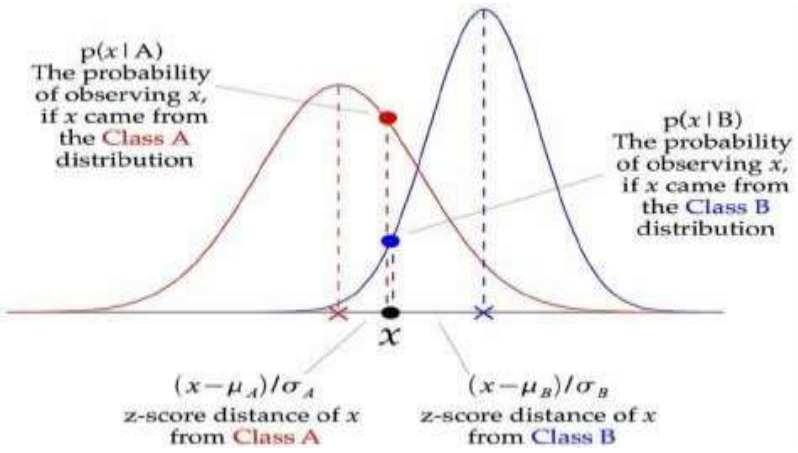
When working with continuous data, it is often assumed that the values associated with each class follow a normal distribution. The likelihood of the features is assumed to adhere to this distribution.

In some cases, it is assumed that the variance is independent of Y (qi), or independent of Xi (σk), or independent of both (σ).

Gaussian Naive Bayes is capable of handling features with continuous values by modelling them according to a Gaussian distribution.

One simple approach to creating a model is to assume that the data can be described by a Gaussian distribution with independent dimensions, meaning there is no covariance between the dimensions. This type of model can be fitted by determining the mean and standard deviation of the data points within each label, which is all that is required to define such a distribution.

Accuracy is 99.09 %

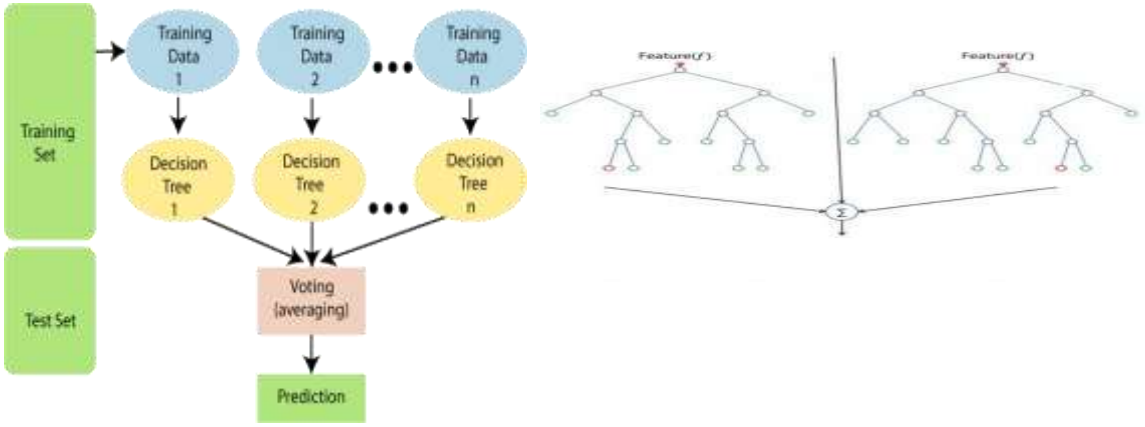


RANDOM FOREST

Random Forest is a supervised learning algorithm that utilizes an ensemble of decision trees, typically trained using the "bagging" technique. Bagging combines multiple learning models to improve overall performance.

For Fertilizer Recommendation, fertilizer data, crop information, and location data are used. This process involves suggesting suitable crops and the corresponding required fertilizer for each crop. External applications are employed to present weather information, temperature data, humidity levels, atmospheric pressure, and overall weather descriptions.

Random Forest serves as a classifier that builds upon decision trees. It comprises numerous decision trees, and when classifying a new instance, each tree offers its own classification for the input data. The Random Forest algorithm aggregates these individual classifications and selects the most frequently predicted class as the final result. Each decision tree in the ensemble is trained using a subset of randomly sampled data from the original dataset. Furthermore, a subset of features is randomly chosen from the available features at each node during the tree's growth process.



XG BOOST

Accuracy is 99.09%

XG Boost is a powerful algorithm that has gained popularity in machine learning hackathons and competitions since its introduction in 2014. It is often praised for its performance and speed in various tasks, such as predicting ad click-through rates and classifying high energy physics events.

The strength of XG Boost lies in its scalability, which enables fast learning through parallel and distributed computing while efficiently utilizing memory. This scalability was recognized by CERN, the European Organization for Nuclear Research, when they were searching for a solution to classify signals from the Large Hadron Collider. The challenge at CERN required a scalable approach to process the massive amount of data generated at a rate of 3 petabytes per year and accurately distinguish rare signals from background noise in complex physical processes. XG Boost emerged as the most useful, straightforward, and robust solution for this task.

To understand how XG Boost works, it's important to know that it is an ensemble learning method. Ensemble learning combines the predictive power of multiple learners to improve overall performance. In some cases, relying on the results of just one machine learning model may not be sufficient. By using ensemble learning, XG Boost offers a systematic solution to harness the strengths of multiple models.

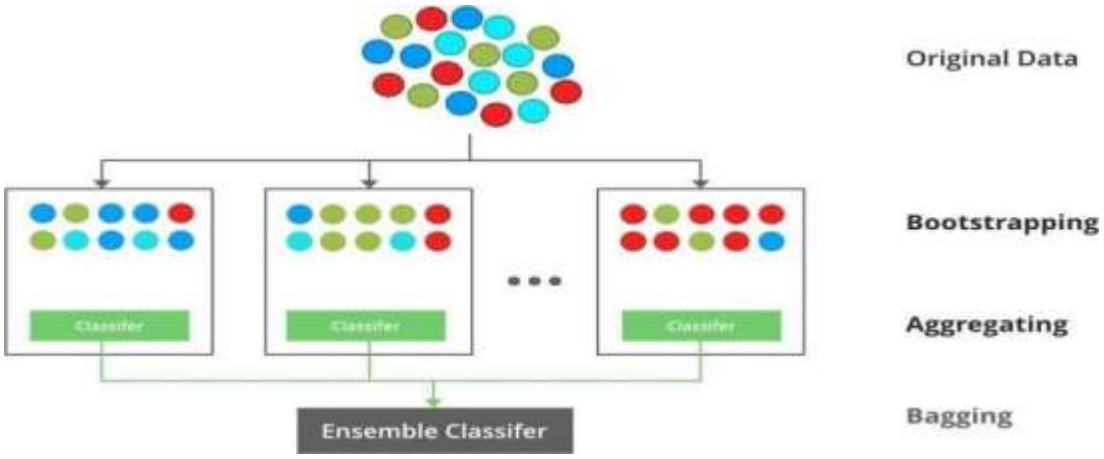
Ensemble learning can be achieved through different techniques, such as bagging and boosting. XG Boost specifically uses a gradient boosting framework and is based on decision trees. Bagging and boosting can be used with various statistical models, but their most prevalent usage has been with decision trees.

Bagging, short for bootstrap aggregating, involves training multiple models independently on different subsets of the training data and then combining their predictions. This helps reduce overfitting and improves the overall stability of the model.

On the other hand, boosting is a technique where models are trained sequentially, and each subsequent model tries to correct the mistakes made by the previous models. In the case of XG Boost, gradient boosting is used, which means that each new model is trained to minimize the errors (residuals) made by the previous model.

The key idea behind XG Boost is the use of gradient optimization techniques to iteratively improve the ensemble of models. It optimizes a specific loss function by minimizing the gradients of the loss function with respect to the predictions of the ensemble. This process is performed in a parallel and distributed manner, making it highly efficient.

In summary, XG Boost is a powerful ensemble learning algorithm that combines the strengths of multiple models, particularly decision trees, using a gradient boosting framework. It leverages parallel and distributed computing to achieve fast learning and efficient memory usage.



PROJECT BACKGROUND

Accurate crop yield prediction is crucial for decision-makers at national and regional levels, enabling rapid and informed decision-making. It helps farmers make informed choices about what and when to grow. Various approaches have been proposed for crop yield prediction. Additionally, the use of mineral fertilizers is essential as they replenish the nutrients that crops extract from the soil, maintaining agricultural productivity.

Another significant aspect is the detection of plant diseases through automated techniques. This approach reduces the burden of manual monitoring in large crop farms by identifying disease symptoms at an early stage.

LIMITATIONS AND CHALLENGES

There are a few limitations and challenges associated with this project:

1. The system requires a network connection.
2. A mobile device is necessary for accessing the application.
3. The project aims to be cross-platform independent.

CONCLUSION

This project highlighted the limitations of current systems and their practical usage on yield prediction. The web application includes multiple features that users can leverage for the selection of a crop. The inbuilt predictor system helps the farmers to predict the yield of a given crop. The inbuilt recommender system allows a user exploration of the possible crops and their yield to make more educated decisions. XG Boost were implemented and tested on the given datasets. Results indicate that XG Boost gives the best result.

REFERENCES

Here are some references that have explored similar topics:

1. S. M. Pandey, P. K. Ramesh, A. Anmol, B. R. Aishwarya, K. Rohilla, and K. Shaurya. "Crop Recommender System Using Machine Learning Approach," 2021.
2. S. Bhanumathi, M. Vineeth, and N. Rohit. "Crop Yield Prediction and Efficient Use of Fertilizers," 2020.
3. C. P. Wickramasinghe, P. L. N. Lakshitha, H. P. H. S. Hemapriya, A. Jayakody, and P. G. N. S. Ranasinghe. "Smart Crop and Fertilizer Prediction System," 2019.
4. Arun Kumar, Naveen Kumar, and Vishal Vats. "Efficient Crop Yield Prediction Using Machine Learning Algorithms," International Research Journal of Engineering and Technology (IRJET), Volume 05, Issue 06, June 2018.
5. Nithin Singh and Saurabh Chaturvedi. "Weather Forecasting Using Machine Learning," 2019 International Conference on Signal Processing and Communication (ICSC).

6. Aakash Parmar and Mithila Sompura. "Rainfall Prediction using Machine Learning," 2017 International Conference on (ICIIECS) at Coimbatore, Volume 3, March 2017.

7. Sachee Nene and Priya R. "Prediction of Crop yield using Machine Learning," International Research Journal of Engineering and Technology (IRJET), Volume 05, Issue 02, February 2018.

8. Ramesh Medar and Anand M. Ambekar. "Sugarcane Crop Prediction Using Supervised Machine Learning," International Journal of Intelligent Systems and Applications, Volume 3, August 2019.