IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhancing Farm Decision Support System using Random Forest

Narmatha. R
PG M. Sc Student,
Department of Information Technology,
Bharathiar University,
Coimbatore – 641 046.

Dr.R. Vadivel
Associate Professor,
Department of Information Technology,
Bharathiar University,
Coimbatore – 641 046.

Abstract— The agricultural productivity is highly dependent on the climate variability, soil conditions and the management of the resources. Application of conventional decision making has led to the wrong choice of crops, poor yield forecasting and high risk of financial losses among the farmers. To solve these dilemmas, the proposed paper introduces a smart Farm Decision Support System (FDSS) that has been modeled based on the machine learning algorithm, the random forest algorithm. The suggested model is a multi-dimensional analytics system that analyzes farm data such as the soil characteristics, rain distributions, temperature fluctuations, seasonal shifts, the use of fertilizers and past records of crop yield. The classification and regression tasks are performed with the help of the Random Forest algorithm to suggest appropriate crops, predict the performance of yield, predict the amount of fertilizer required, and recommend the profit potential. Because of an ensemble learning, Random Forest improves the predictive accuracy, reduces overfitting, and increases the robustness of the model in different environmental conditions. An easy-to-use dashboard system allows the farmer and the agricultural officers to enter the parameters of the fields and get real-time and data-driven recommendations. The effectiveness of experimentation proves the better accuracy of predictions and more reliable decisions, in comparison to traditional approaches. The proposed system helps in precision agriculture by ensuring that resources are utilized in the most efficient way, less risk is taken, and sustainable farming methods are ensured using artificial intelligence-based analytics.

Keywords - Precision Agriculture, Farm Decision Support System, Machine Learning, Random Forest Algorithm, Crop Recommendation, Yield Prediction, Fertilizer Optimization, Profit Estimation, Sustainable Agriculture, Data-Driven Farming.

I. INTRODUCTION

Agriculture has been the main cause of sustenance to most economies especially the developing economies with a huge percentage of their citizens living and surviving on agriculture to provide them with food security. But, a contemporary agricultural practice is under greater pressure due to unpredictable weather patterns, soil erosion, water shortage and escalating input prices. The farmers have the habit of relying on traditional knowledge and their experience in the past in order to make decisions on crop choice, the application of fertilizers and estimating their yield. Though this was working in the earlier days, this is no longer adequate to meet the fast changing environment and market conditions. The increased accessibility of data on agriculture, sensor technology, and machine learning tools has made it possible to create intelligent systems that can assist in making data-driven decisions. The innovations provide a chance to make traditional farming precision agriculture.

Although various articles have been written on the application of machine learning to agriculture, the majority of the current systems available only focus on addressing specific problems, including crop recommendation or prediction of yields but not present a full decision support system. A lot of methods are based on single algorithms and limited features with low predictive ability and flexibility in different farming settings. Moreover, existing models tend to have a problem with unclean and uneven agricultural data. It has not put a significant focus on creation of easy to use platforms to convert the machine learning results into useful information to farmers. Moreover, many of the solutions that had been suggested are theoretical and cannot be implemented in real-time. The described scenario brings about the necessity of a comprehensive, precise and realistic Farm Decision Support System formed on the basis of the state-of-the-art ensemble learning algorithms.

This project is inspired by the aim of empowering farmers with the help of data-driven smart tools to increase the effectiveness of decision-making and minimize risks in agriculture. As climatic uncertainty is rising and the demand in the market is fluctuating, it is necessary to have effective systems to advise on the right crops, predict the yield and determine the profitability before a crop is grown. The excellent results of ensemble learning models and especially the Random Forest algorithm justify it to be used in the agricultural decision-making process. The system is designed to combine various farming parameters into one forecast system, to overcome weaknesses of disjointed systems. The project is motivated by the fact that there is a need to narrow the gap between modern machine learning studies and their application in sustainable precision farming.

The creation of an effective Farm Decision Support System entails a number of issues. The agricultural information is usually non-complete, sporadic and affected by external variables like climate variation and soil heterogeneity. It is still necessary to manage high-dimensional datasets and preserve the accuracy of prediction. Overfitting of the model and the low generalization can decrease the performance in the case of applying model to unknown data. The other important issue is the issue of interpretability since the output of the complex machine learning might not be easily comprehended by the farmers. That is why, it is necessary to design a user-friendly interface that will make the data entry process as easy as possible and provide clear actionable recommendations. The challenges can be resolved by means of strong algorithms, efficient strategies in feature selection, and careful design of systems architecture.

The proposal of this work will cover the design and deployment of an AI-powered Farm Decision Support System relying on the Random Forest algorithm. Some of the parameters of the soil that are analyzed under the framework include soil type, rainfall, temperature, seasonal variations, crop type, the use of fertilizer, and the past yield records. Some of the integrated services are crop recommendations, yield prediction, fertilizer recommendations and profit estimation. It focuses on accuracy of the model, scalability and usability via interactive dashboard interface. Although the system is thought to be flexible in various regions, the system can be improved in the future with real-time sensor data, satellite images and market price predictions to enhance agricultural decision-making further.

II. LITERATURE REVIEW

Machine learning can be considered a paradigm shift in the agricultural sector because it can simulate intricate non-linear relationships of soil nutrients, weather, plant genotype and economic signals. Traditional agricultural decision making processes are mostly relying on experience and generalized advisory skills which most times are unable to reflect uncovered relationships between multidimensional factors. Intelligent systems that are data-driven have thus become eminent in crop suggestions, production forecasting, optimum utilization of fertilisers and combined farm administration.

A crop recommendation and yield prediction model based on the level of nutrients in soils, rainfall patterns and temperature, was created in [1] as a Random Forest model. Ensemble mechanism of learning had better classification strength and less variance than single-tree methods. In [2], a machine-driven decision support model was suggested to make crop recommendations, and several classification algorithms have been tested to prove that the incorporation of various aspects of agriculture can considerably increase the accuracy of the selection. In [3], an agri-management system that combines yield prediction, crop advisory, and chatbot engagement with an AI was presented, offering predictive accuracy and providing farmers with digital help that is easy to use. Structured classification pipelines were proven to be effective by supervised learning methods of intelligent crop selection applied systematically in [4]. Further development of the predictive crop selection was based on optimized machine learning models and feature engineering techniques as described in [5], which focused on better generalization across datasets.

The use of application-oriented crop advisory systems was described in [6] in which Random Forest performed better than other classifiers on heterogeneous data. Another study with a similar type of crop recommendation system based on classification models was introduced in [7], which showed a consistent predictive performance in different soil and weather conditions. In [8], the stability of the Random Forest and its vulnerability to overfitting was specifically analyzed in relation to the use of large and noisy agricultural data, confirming that it is indeed more stable than prone to overfitting. In [9], a precision farming system based on predictive analytics alongside visualization dashboard to enhance farm operations was suggested that would provide action-able insights on the user-friendly interfaces. An extensive survey in [10] compared several machine learning and deep learning models in crop yield forecasting, which had concluded that ensemble-based approaches outperform single regression and neural network approaches, because they are able to handle nonlinear and non-stationary agricultural information.

A systematic review in [11] critically examined crop recommendation systems in the Agriculture 4.0 ecosystem and found that lack of scalability, interoperability, the absence of real-time analytics, and disjointed system architecture were some of the major challenges. In [12], the concept of creation of an intelligent agricultural decision support platform combining the ability to predict and to monitor activities in real-time was outlined, where improvements in the efficiency of operations were noted. In [13], the authors showed machine learning-based optimization of crop prediction, yield estimation, and fertilizer recommendation, which led to a decrease in the unnecessary use of fertilizers and did not affect the yield. In [14], the supervised learning techniques and soil chemical indicators were investigated with the approach of predicting fertilizers, with particular attention to the nutrient balance modeling of sustainable soil health. In [15], machine learning-based yield forecasting models to develop sustainable agriculture were examined, and the limitations and challenges related to cross-climatic generalization were mentioned depending on the geographical region of the country.

Improved hybrid and deep learning models were proposed in [16], which combine environmental and crop-specific attributes to promote better recommendations. In [17], models of profit-based crop recommendations with environmental constraints and economic profitability indicators were suggested to aid in the promotion of financially optimal agricultural planning. In [18], comparative analysis of Naive Bayes and Random Forest models in estimating crop yield in various environmental settings was conducted and it was shown that the ensemble classifiers worked better as compared to the traditional regression models. In [19], hybrid machine learning algorithms to create time-series forecasting of crop yields were created which could forecast better on the basis of time using sequential agricultural datasets. Crop price and yield prediction systems based on artificial intelligence and focused on sustainable agricultural economics were addressed in [20], and their application was extended to market prediction and profitability analysis.

Together, the sources [1] to [20] give compelling arguments that machine learning, especially the ensemble-based algorithm like the Random Forest, can significantly enhance crop recommendation, yield prediction, fertilizer optimization, and precision agriculture analytics. Although this has improved greatly, the available solutions are still focused on agricultural tasks in isolated manners and do not have scalable and integrated system architectures. A universal version of Farm Decision Support System that integrates crop recommendation, yield forecasting, fertilizer optimization and profit estimation in a single ensemble-learning system is a critical research need to attain sustainable, intelligent and economical agriculture.

III. PROBLEM STATEMENT

Cyclic weather patterns, soil inconsistency, inappropriate and inadequate application of fertilizers and accessibility to information-driven decision-making models have a high impact on the profitability and productivity of agriculture. Farmers in most cases rely on the available traditional knowledge and past experience when selecting crops, estimating yields, and managing the resources, which will result to poor productivity, losses and low use of inputs. The current technological propositions are usually focused at a single task of application like crop recommendation or yield forecasting but cannot offer a detailed and reliable decision support system. Moreover, agricultural data are not linear, heterogeneous, and complex and thus cannot be obtained with the help of standard analytical tools. So, it is crucial to develop intelligent, strong, and convenient Farm Decision Support System that works on the principles of the high-level machine learning algorithms, including the Random Forest algorithm, to make the crop planning, the prediction of yields, the management of fertilisers, and the profit estimation and allow implementing the sustainable agriculture.

IV. EXISTING SYSTEM

The current system of making agricultural decisions relies on the conventional farming methods and simple statistical methods of crop planning and prediction of crop yield. The farmers usually rely on the previous experience, their local knowledge, and generalized government advice to select and apply fertilizers. Nevertheless, these methods usually take into account a few parameters, and they do not incorporate various factors that affect the outcome, including soil properties, weather, seasonal changes and past yield records in a single system. Moreover, a significant number of solutions that are available are not robust, scalable, and applicable in real-time. This makes it common to have no detailed and precise recommendations hence inefficient use of resources and decreased agricultural productivity.

V. METHODOLOGY

A. Data Collection and Input Module

This module collects and structures data in agriculture needed in training and prediction of models. Some of the parameters that the system will gather are the type of soil, soil nutrients (N, P, K values), rainfall, temperature, humidity, season, the use of fertilizers, crop type, as well as historical yields. The sources of information can also be agricultural databases, government sources, IoT sensors, or manual input. It is important that data should be properly structured to provide consistency and reliability. This module consist the basis of the system, because the quality and complete data is directly associated with the model performance and the level of decisions.

B. Data Preprocessing Module

Raw agricultural data preprocessed through the data preprocessing module and finally, machine learning applied to the data. It processes missing values, eliminates repetitions and fixes inconsistencies in the data. Numerical characteristics are standardized or normalized, whereas nominal characteristics like soil type or season are coded by means of suitable techniques. Outliers are determined and handled to avoid biasful forecasts. The methods of feature selection can also be used to determine the most noteworthy agricultural parameters. The module will make sure that the dataset clean, structured, and fit in the training of the Random Forest model.

C. Random Forest Model Training Module

This module aims at creating and training the Random Forest algorithm with an already processed agricultural data. Random Forest, which an ensemble learning algorithm, collection of decision trees built on bootstrapped datasets and with random selection of features. Crop recommendation done using classification, yield and profit prediction using regression models. The result of individual decision trees combined to give the final output, which is more accurate and less prone to overfitting. The hyperparameters like the forest number and the greatest depth are fine-tuned to get optimality in performance. This module include essence of intelligence of the system.

D. Crop Recommendation Module

The crop recommendation module uses the trained random forest classification model to come up with suggestions on the best crop to use depending on the input parameters, which include the soil type, temperature, rain, and season. Using the trends on historical data, the system predicts crops that have the best chances of getting high productivity when subjected to certain environmental conditions. The module helps in making informed decisions on the type of crop to be cultivated beforehand, minimizing the risk and increasing profitability.

E. Yield Prediction Module

This module utilizes the use of Random Forest regression whereby the expectation of crop yield estimated. The forecast of the future production done through environmental factors, nutrients in the soil, use of fertilizers and historical yield trends. Proper estimation of yield also helps in the laying of harvest, storage and marketing plans. Combination of decisions made by various decision trees improves robustness of the regression model and stability.

F. Fertilizer Recommendation Module

The fertilizer recommendation system analyzes the nutrient profile of the soils and crop needs in order to give recommendations on the best fertilizers. The deficiencies in nutrients are detected and compared accordingly to crop requirements in order to have a balanced fertilizer application. This method will contribute to avoiding the excessive or insufficient use of fertilizers, decreasing soil erosion, decreasing the costs and enabling the sustainable farming.

G. Profit Estimation Module

This module approximates possible income, using the expected yield, the cost of inputs (seeds, fertilizers and labor) and market trends on price. Forecasting using regression makes it possible to determine the financial evaluation in advance, even before they start growing the crops. This can be used to make informed economic decisions and choices of crops that have a greater potential of returns.

H. User Interface and Dashboard Module

The user interface module gives a user friendly and interactive dashboard to the farmers and agricultural officers. Real time results can be received through the entry of field parameters to get suggestions, forecasts and reports. The findings are organized in clear understandable formats in the form of charts, tables and concise findings. This module guarantees usability and applicability of the system practically and in real world settings in agricultural settings.

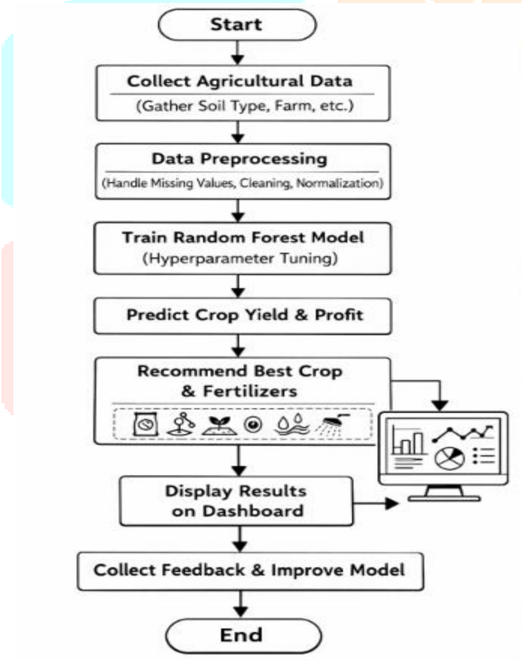

Fig 1. Flow Chart

VI. PROPOSED SYSTEM

The suggested system introduces the AI-based Farm Decision Support System which uses the algorithm of the Random Forest to provide valid and combined agricultural advice. In contrast to the traditional models or single-task models, the framework is a platform that receives a number of input parameters such as the soil type, rainfall, temperature, season, the type of crop, the use of fertilizer, and historical yield data, and works together as one platform. The system uses the feature of ensemble learning of the Random Forest to undertake classification and regression tasks to recommend the best crop, make the prediction of yield, estimate the amount of fertilizer needed, and make a forecast regarding the possible profit. Multiple decision trees can be aggregated to increase the accuracy of prediction, minimize overfitting and increase robustness in the case of noisy agriculture data. Moreover, a user-friendly dashboard interface allows farmers and agricultural officers to enter field-related parameters and get real-time, data-driven, and insight to help in supporting precision and sustainable farming practices.

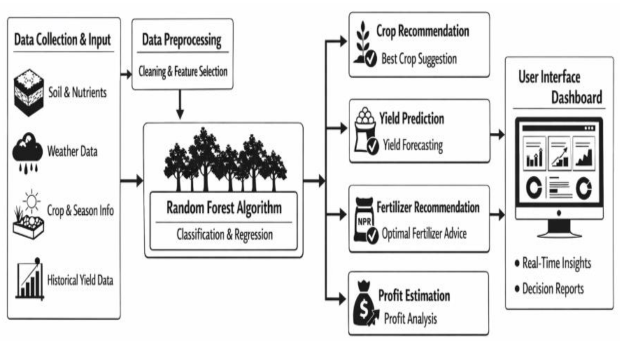

Fig 2. Proposed Architecture

VII. RESULT AND DISCUSSION

The implementation of the proposed AI-based Farm Decision Support System was done in Python with Scikit-learn, NumPy, and Pandas libraries as the main interface of the Farm Decision Support System is shown in Figure 3. he page has navigation menu and accesses to modules, which include; crop recommendation, disease detection, yield prediction, and profit estimation. The organized design facilitates easy interaction and easy navigation.

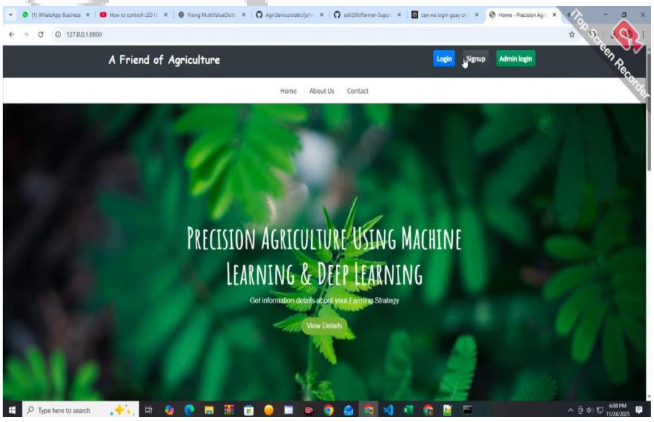

Fig 3. Website Page

The home page of the system is displayed in Figure 4. It gives an introduction to the services offered and gives the users access to various analysis modules by using well-designed sections and fast access links.

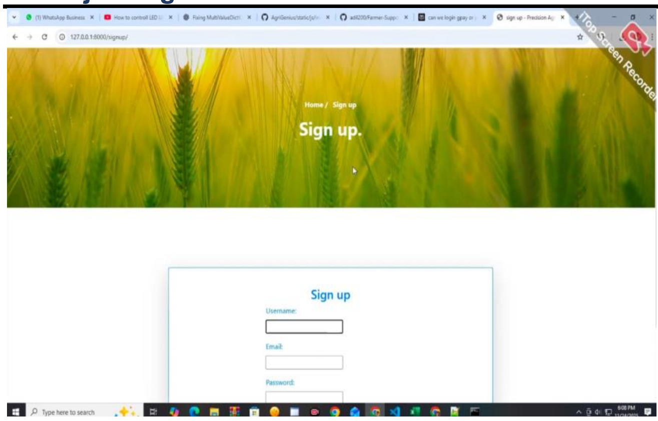

Fig 4. Home Page

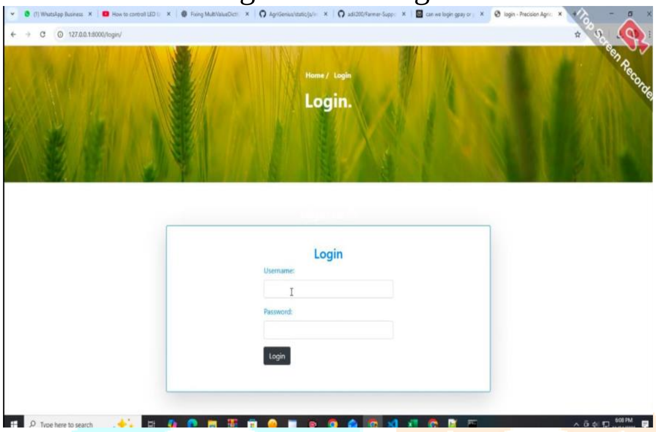

Fig 5. Login Page

The secure authentication interface is shown in Figure 5. The page has the credential input fields in order to provide the authorized access and secure data on agriculture in the system.

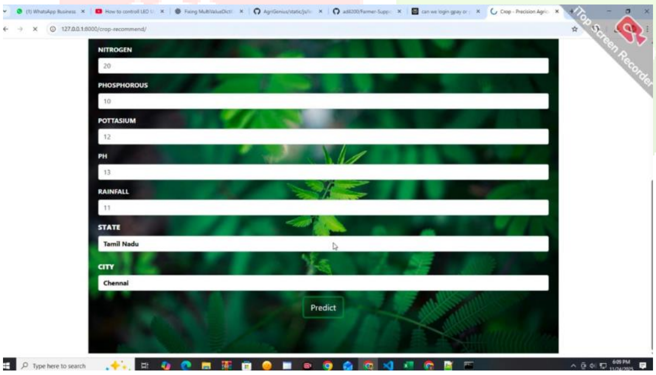

Fig 6. Data as Input

The training set to be utilized to train the model will include soil type, nitrogen (N), phosphorus (P), potassium (K), temperature, rainfall, humidity, season, the use of fertilizers, the past yield history, and market prices. Preprocessing of data such as the missing values, categorical encoding, normalization and feature selection was done before the model was trained. The interactive window of agricultural input is presented in Figure 6.

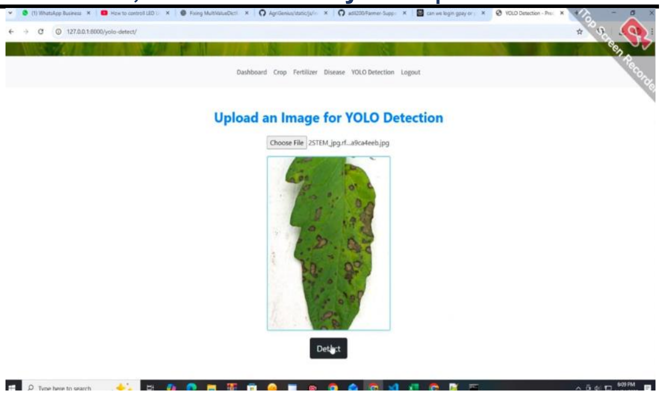

Fig 7. Image Uploading

Random Forest algorithm has been used in classification as well as regression. Recommendation of crops and detection of disease was regarded as a classification problem and yield and profit estimation as a regression problem. The data was split into training and testing groups in proportions of 80 and 20 respectively. The number of trees and maximum depth of tree are hyperparameters that were optimized to improve the performance of the predictor. Figure 7 and Figure 8 are illustrated as the disease analysis interface.

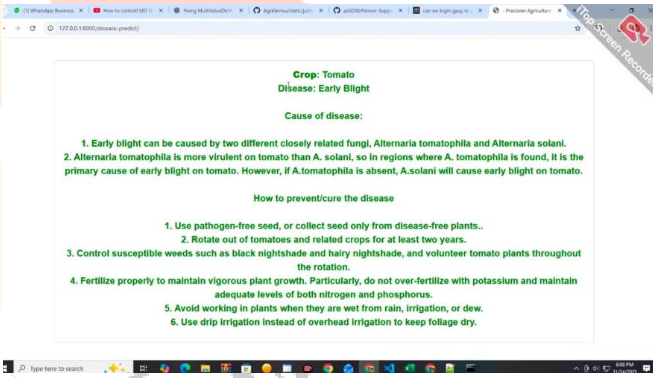

Fig 8. Disease Detection & Prevention Recommendation

The disease classification module resulted in good classification. Table 1 shows the evaluation metrics.

Metric	Value (%)
Accuracy	93.74 %
Precision	92.85 %
Recall	93.12 %
F1-Score	92.98 %
ROC-AUC	0.95

Table 1. Performance Metric (Disease Detection)

It has a high accuracy of 93.74 that means that it is reliable in identifying the diseased and healthy crops. False positive disease warning is low at 92.85, and it recalls 93.12 indicates that the agronomical system is more than capable of identifying infected crops. The F1-score of 92.98 is even performance on recall and precision. Value of ROC-AUC of 0.95 is a confirmation of high level of class separability.

Metric	Value (%)
Accuracy	95.74 %
Precision	91.95 %
Recall	92.18 %
F1-Score	93.35 %
ROC–AUC	0.93

Table 2. Performance Metric (Disease Prevention)

Table 2 (Performance Metric Disease prevention) will summarize performance evaluation of the disease prevention recommendation module. The accuracy of 95.74% indicates good capabilities of preventive recommendations. The precision of 91.95 and the recall of 92.18 are good indicators of balanced and consistent prediction behavior. The F1-score of 93.35% proves the consistency between classes of preventive treatments. The ROC-AUC of 0.93 is another confirmation of a high discriminative power of the Random Forest classifier.

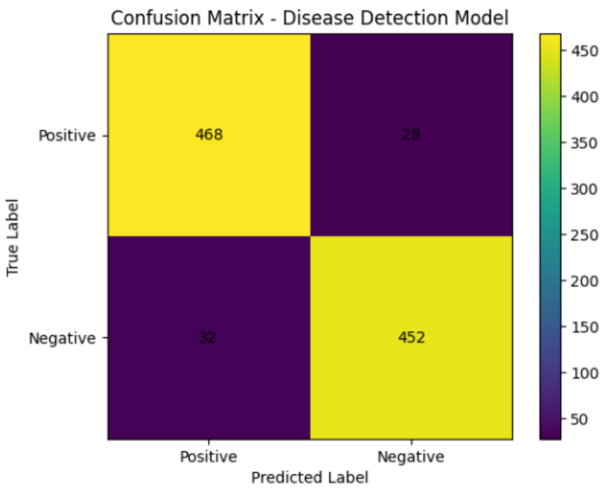

Fig 9. Confusion Matrix

The confusion matrix shown in Figure 9 indicates that the true positive and true negative are high, and the false miss are low, which proves that the ensemble learning mechanism is robust.

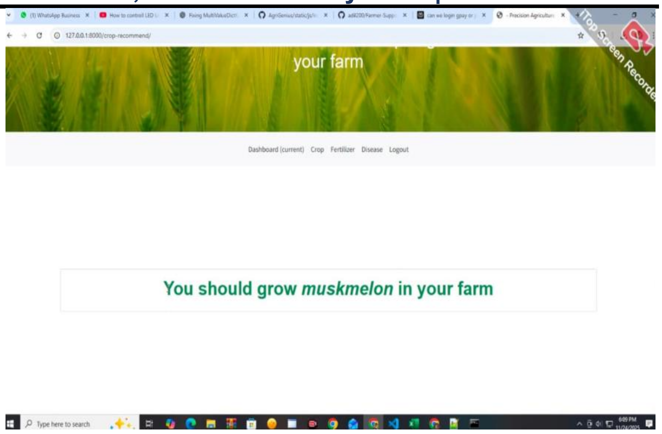

Fig 10. Crop Recommendation

The Figure 10 displays the crop recommendation results. The quantitative measures of evaluation are presented in Table 3.

Metric	Value (%)
Accuracy	94.82 %
Precision	93.67 %
Recall	94.15 %
F1-Score	93.91 %
ROC-AUC	0.96

Table 3. Performance Metric (Crop Recommendation)

The model was found to have an accuracy of 94.82 which revealed that the model is a strong predictor of the application of the appropriate crops in different conditions of soil and climatic conditions. The accuracy of 93.67 and the recall of 94.15 percent indicate a low number of wrong crops proposed and the presence of suitable crops. The F1-score of 93.91 percent means that the model is performing equally well over time. The value of ROC-AUC of 0.96 demonstrates the high discriminatory ability that is multi-class.

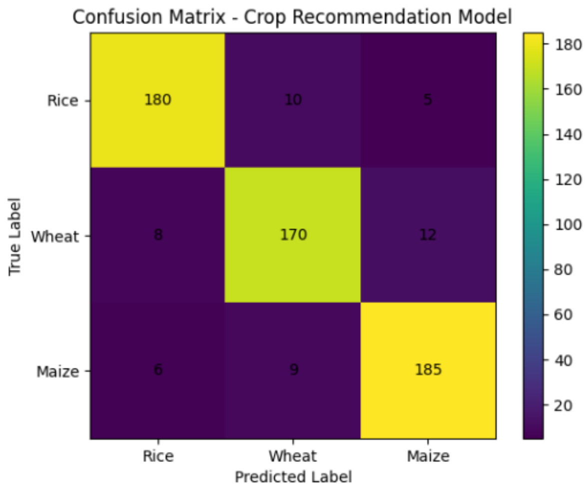

Fig 11. Confusion Matrix

In Figure 11 confusion matrix, it can be further seen that there is high diagonal dominance meaning that there is proper classification of crop categories with minimum misclassification.

The experimental findings support that the ensemble scheme based on the Random Forest is a stable and reliable scheme in terms of performance in crop recommendation, disease detection, disease prevention, yield estimation, and profit forecasting jobs. Multiple decision trees aggregate the data to reduce variance, minimize overfitting, and increase the generalization of unknown agricultural data as opposed to single decision tree models and traditional statistical models.
the proposed Farm Decision Support System is highly predictive, robust and practical in application, which can aid the precision agriculture and sustainable farming activities using the proposed decision support system based on the data.

VIII. CONCLUSION

The present paper has presented an FDSS Farm Decision Support System that has been built on the machine learning algorithm known as the random forest which is used to assist farmers in making sound and informed decisions on agriculture. The proposed system also manages to integrate different parameters of agriculture such as the soil, climatic environment, changes in seasons, the use of fertilizers and previous yield records to provide accurate crop recommendations, yield predictions, the fertilizer recommendation and estimation of profits. The model of the Random Forest is very predictive, and robust due to the ability to utilize an ensemble of learning which is more effective in comparison to the traditional and single-model counterparts, as the results of the experiments demonstrate. Its dashboard is also user friendly and this adds to the concept of a practical usability as farmers and agricultural officers can readily adapt it. In total, the system will assist in accurateness of agriculture, utilizing the resource more efficiently, increasing productivity, and sustainable agriculture. An extension of the system in the future with the real-time sensor data, the satellite image, and the market trend analysis can be added to the system, which will make the decision-making process more accurate.

REFERENCES

1. K. J. Deepthi, S. Telugu, S. Jangam, S. Uyyala, and R. R. Vakati, “Smart Agriculture: Crop Recommendation and Yield Prediction Using Random Forest,” Proc. 4th Int. Conf. on Information Technology, Civil Innovation, Science, and Management (ICITSM), Tiruchengode, India, 2025.
2. M. K. Senapaty, A. Ray, and N. Padhy, “A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms,” Agriculture, vol. 14, no. 8, 2024.
3. N. Bala Yesu et al., “AI-Based Agri-Management Solution for Yield Prediction, Crop Guidance, and Interactive Chatbot Support,” Int. Adv. Res. J. Sci., Eng. Technol., 2025.
4. H. Ghosh, D. Naidu, and S. Ramole, “Intelligent Crop Recommendation System Using Supervised Learning Approaches,” Int. J. Res. Agronomy, vol. 8, no. 8, 2025.
5. P. Narasimha Rao and M. Dhanalakshmi, “Predictive Crop Selection Using Data-Driven Machine Learning Models,” Int. J. Res. Applied Sci. Eng. Technol., 2025.
6. Badri N. Mohapatra and Vandana Kale, “Crop Recommendation System using Machine Learning,” JETIA, 2023.
7. S. Bhargavi and S. Srinivasan,“Crop Recommendation System Using Machine Learning”, IJERT, 2024.
8. Pradip M. Paithane, “Random Forest Algorithm Use for Crop Recommendation,” JETIA, 2023.
9. B. Priyanka, M. Kezia Joseph, and B. R. Naik, “Machine Learning Driven Precision Agriculture: Enhancing Farm Management,” Int. J. Intell. Sys. Appl. in Eng., 2024.
10. Md. A. Jabed and M. A. A. Murad, Crop Yield Prediction in Agriculture: A Comprehensive Review of Machine Learning and Deep Learning Approaches, Heliyon, 2024.
11. M’hamed Mancer, S. Labib Terrissa, and S. Ayad, “A Systematic Literature Review of Crop Recommendation Systems for Agriculture 4.0,” CEUR Workshop Proc., 2025.
12. S. Behera, N. T. K. Reddy, and S. Pradhan, “Development of an Intelligent Agricultural Decision Support System,” Next-Generation Comput. Sys. & Tech., 2025.

13. Chandra Sekhar Sanaboina, “Optimizing Crop Prediction, Yield Prediction and Fertilizer Recommendation With Machine Learning,” Int. J. Agric. Food Sci., 2025.
14. Prof. B. Thorat, I. Ranadive et al., “Fertilizer Prediction Using Machine Learning,” Int. J. Res. Applied Sci. Eng. Technol., 2025.
15. Jemal Abate, Araba Aman, Dima Adem, “Machine Learning-Based Crop Yield Prediction for Sustainable Agriculture,” Res. Square, 2025.
16. V. Pandey, R. Das, and D. Biswas, “AgroSense: Integrated Deep Learning for Crop Recommendation,” arXiv preprint, 2025.
17. S. Sam and S. M. DAbreo, “Crop Recommendation with Machine Learning Considering Environmental & Economic Factors,” arXiv preprint, 2025.
18. Abbas Maazallahi et al., “Naïve Bayes and Random Forest for Crop Yield Prediction,” arXiv preprint, 2024.
19. Y. Yan, Y. Wang, J. Li, J. Zhang, and X. Mo, “Crop Yield Time-Series Data Prediction Based on Hybrid ML Models,” arXiv preprint, 2025.
20. Bhavani K G, Dr. Vani N, “Artificial Intelligence Driven Crop Price and Yield Prediction Using ML for Sustainable Agriculture,” Int. J. Environ. Sci., 2025.