Machine Learning-Powered Agricultural Crop Prediction with CatBoost and Web Interface

Chanikya.A
221FA04166
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Guntur, India

Mohan koteswar .T
221FA04215
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Guntur, India

Hari krishna sri.B
221FA04607
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Guntur, India

Swathi.k
221FA04700
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Guntur, India

Abstract: Agricultural productivity is significantly influenced by the choice of crops suitable for the local soil and environmental conditions. The "Smart Crop Advisor" uses machine learning to assist farmers in selecting the best crop for cultivation based on real-time environmental and soil parameters. This research uses a dataset containing various attributes related to the soil and weather conditions of the farmer's field. The dataset is pre-processed and scaled using Standard Scalar normalization, and multiple classification algorithms are trained and evaluated. The Random Forest emerged as the most accurate model for predicting crop suitability with 99.7%, followed by CatBoost and Naive Bayes with 99.2% and 99%, respectively.The proposed solution accepts user input for environmental parameters, scales the data appropriately, and predicts the most suitable crop using the trained model. This system provides a practical and efficient tool for farmers to make informed decisions about crop cultivation, enhancing yield, optimizing resource usage, and contributing to sustainable agriculture practices.

Keywords: Smart Crop Advisor, Machine Learning, Data Scaling, Random Forest Classifier, Predictive Modeling

I. INTRODUCTION

Most Agriculture is a crucial part of economies around the world, playing a key role in ensuring food security, supporting rural communities, and promoting sustainable development. However, farming today faces many challenges, including unpredictable weather, soil degradation, water shortages, and pest infestations. These problems are made worse by climate change, making traditional farming methods less reliable and effective. Farmers need to make important decisions about which crops to grow to get the best yield and profit while minimizing environmental impact. Traditionally, these decisions have been based on experience, intuition, or general guidelines that may not consider the specific conditions of a particular location, often leading to less-than-ideal results.

Advanced technologies and data science offer new solutions to these challenges, and machine learning (ML) has become a powerful tool for improving decision-making in agriculture. A "Smart Crop Advisor" system uses machine learning algorithms to analyze various types of data, such as soil properties, weather forecasts, past crop performance, and market trends, to suggest the best crops for a specific location and time. These systems can adjust to local conditions, providing recommendations that are more accurate than traditional methods. Machine learning models have shown great potential in this area because they can learn patterns from large amounts of data. Techniques such as Gradient Boosting Machines (GBM), CatBoost, Linear Discriminant Analysis (LDA), and Quadratic Discriminant Analysis (QDA), and Ensemble Learning have been successfully used in agriculture for tasks like predicting yields, detecting crop diseases, and modeling climate. For crop recommendations, these models can analyze factors like soil pH, nutrient levels (such as nitrogen, phosphorus, and potassium), temperature, rainfall, and humidity to predict which crops will perform best in certain conditions.Research has shown that machine learning models can achieve high levels of accuracy in recommending crops. For example, methods that combine multiple algorithms, known as ensemble learning, have been found to reach accuracies over 95%. These advancements highlight the potential of machine learning to improve crop selection by providing reliable, data-driven details related to local environments. Smart crop recommendation systems can help farmers maximize yields, reduce input costs, and promote sustainable farming practices by making better use of resources like water, fertilizer, and labor.

This research aims to develop a strong crop recommendation model using the latest machine-learning techniques. The study will use a detailed dataset that includes different environmental and farming variables from various regions to build a model that is both flexible and applicable in different contexts. The research will focus on selecting the most important variables, fine-tuning the model to improve performance, and validating it with independent datasets to ensure it works well in real-world situations.

By examining different machine learning methods and how they can be applied to crop recommendations, this paper contributes to the growing field of precision agriculture. The ultimate goal is to develop a smart crop recommendation system that helps farmers make informed decisions, reduce uncertainty in crop planning, and support sustainable agricultural practices that can adapt to changing climate and economic conditions.

II. RELATED WORK

The research [1] explores the use of machine learning to improve agricultural decision-making in India, focusing on crop selection and yield prediction to help farmers make informed choices. The study developed two modules: one for predicting suitable crops and another for estimating yields. For crop prediction, various algorithms were tested. The Naive Bayes classifier achieved the highest accuracy with a Cohen's Kappa score of 95.13%. For yield prediction, models including Multiple Linear Regression, Random Forest Regression, Support Vector Regression, and KNN Regression were evaluated. The Random Forest Regressor performed best, with R-squared values of 82.85% and 91.79% across two datasets, indicating strong predictive capability. The researchers also developed a web application using Flask, enabling farmers to input specific conditions to receive tailored crop recommendations and yield predictions.

Paper [2] investigates the use of ML for crop prediction, which is vital for agricultural planning. It compares three algorithms—K-Nearest Neighbor (KNN), Decision Tree, and Random Forest Classifier—using the crop recommendation dataset. The dataset was divided into 80% training and 20% testing data, and the models were assessed based on various metrics. Results showed that the Random Forest Classifier performed best, achieving a 99.32% accuracy, followed by the Decision Tree Classifier with 98.86% (Gini criterion) and 97.95% (Entropy criterion). KNN had the lowest accuracy at 97.04%. The study highlights the potential of machine learning in aiding farmers with crop selection based on environmental factors.

The study [3] presents a Machine Learning-based Crop Recommendation System that addresses challenges like unpredictable climate, diverse soil types, and limited modern farming access. The researchers evaluated nine machine learning algorithms, including Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, and Random Forest, using historical data on climate, soil properties, and crop yields.The Random Forest algorithm achieved the highest accuracy at 99.31%. The study highlights the system's potential to improve crop productivity, manage risks from climate and market changes, and promote sustainable practices. While recognizing limitations in capturing land quality and climate variations, the authors suggest future research to address these gaps, contributing to global food security and precision agriculture.

The paper [4] introduces an AI-enabled crop recommendation system that uses ML to suggest optimal crops based on soil composition and weather patterns, aiming to help farmers maximize yields and profitability. Using a dataset from Kaggle containing soil parameters and climate factors, the researchers tested seven machine learning models, including Decision Tree, Naive Bayes, Random Forest, k-nearest Neighbors (k-NN), Support Vector Machine (SVM), XGBoost, and Logistic Regression. Naive Bayes and XGBoost achieved the highest accuracy of 99.55%, while Random Forest and SVM followed with 99.31%. The system's high accuracy suggests it could effectively support precision agriculture by optimizing resource use, reducing crop failures, and boosting productivity. The study also discusses potential enhancements, like integrating IoT and remote sensing technologies, and suggests future research involving diverse datasets and deep learning methods.

Research [5] explores the use of ML for crop recommendation in smart farming the authors tested seven algorithms, including Logistic Regression, Decision Trees, Random Forests, K-Nearest Neighbors, Naive Bayes, Support Vector Machines, and Neural Networks, using a Kaggle dataset with various soil and weather features. The Random Forest algorithm achieved the highest accuracy at 99.5%, while the Neural Network reached 97.73% with 1000 epochs. The study highlights the potential of these models to improve crop yields, sustainability, and profitability, and discusses challenges like climate change, water scarcity, and data quality. Future directions include assessing the economic impact through farmer surveys, developing mobile apps, and incorporating real-time sensor data.

The research paper [6] introduces a new approach to crop recommendation using ML through a hybrid model called the Wrapper-PART-Grid method. This model combines a wrapper feature selection method to identify key soil features, grid search (GS) for hyperparameter optimization. Using a dataset of 2,200 instances with various soil and weather parameters, the model achieved the highest accuracy of 99.31%, outperforming other ML models such as multilayer perceptron, instance-based learning, and decision trees.

The paper [7] presents a system that combines machine learning (ML) and deep learning (DL) models to improve crop recommendations for farmers. Utilizing a crop recommendation dataset available in Kaggle, the study evaluated models including Decision Trees, Random Forests, XGBoost, SVM, KNN, Naive Bayes, ANN, DNN, and Temporal Convolutional Networks (TCN). The TCN model achieved the highest accuracy of 99.9% by effectively capturing temporal dependencies, while the Random Forest model also performed well with 99.2% accuracy. The study emphasized parameter optimization for XGBoost and TCN, highlighting their precision. A web-based interface was developed to allow farmers to input soil and environmental data and receive immediate crop recommendations.

III. BACKGROUND

A Weighted Agriculture plays a crucial role in global food production and economic stability. However, traditional farming methods face significant challenges due to climate variability, soil degradation, water scarcity, and inefficient crop selection. Farmers often rely on experience-based decision-making, which lacks precision and does not account for real-time environmental changes. The integration of machine learning (ML) in precision agriculture has emerged as a transformative solution, enabling data-driven crop recommendations based on soil and weather conditions. Machine learning models excel in recognizing complex patterns within large agricultural datasets, providing highly accurate and adaptive crop recommendations. The Kaggle Crop Recommendation dataset, which consists of essential soil and climate parameters (N, P, K levels, temperature, humidity, pH, and rainfall), serves as a foundational dataset for training predictive models. Supervised learning algorithms such as Random Forest, Naïve Bayes, and Gradient Boosting Machines (GBM) have demonstrated high accuracy in classifying crops based on soil properties and environmental conditions. These models help in optimizing resource utilization, improving yield predictions, and promoting sustainable farming practices.

Several studies have explored machine learning-based crop prediction, achieving high accuracies through ensemble learning, decision trees, and probabilistic classifiers. However, existing systems often lack real-time integration, deep learning enhancements, and accessibility via mobile or IoT-based solutions. This research aims to address these limitations by developing a web-based Smart Crop Advisor that leverages multiple ML models, data preprocessing techniques, and hyperparameter tuning to deliver high-precision crop recommendations. The proposed system not only improves decision-making for farmers but also contributes to the advancement of AI-driven precision agriculture.

IV. PROPOSED CROP RECOMMENDATION METHODS

A. Data Collection and Preprocessing

The foundational step in developing the ”Smart Crop Advisor” is the collection of relevant data that encompasses both environmental and soil parameters crucial for crop growth. The flow of the proposed method is shown in Fig 2 and Fig 3. Data preprocessing is a critical step in preparing the raw dataset for ML analysis.Removing of outliers from numeric data is a crutial step in ML as outliers effect the performance of the model being trained. So, to remove outliers, all the numeric columns are selected first, then the first quartile (Q1) and third quartile (Q3) for each of these columns is calculated. The interquartile range (IQR) is computed as the difference between Q3 and Q1, representing the middle 50% of the data. This process identifies potential outliers as any data point that falls below Q1 minus 1.5 times the IQR or above Q3 plus 1.5 times the IQR. Finally, the rows containing these outliers are filtered out, resulting in a cleaned DataFrame where the outlier values have been removed from the numeric columns. After removing outliers, the dataset is divided into features and target variables. Features include the environmental and soil attributes, while the target variable is the crop label.Following this, the data is split into training and testing sets, with 80% of the data allocated for training the models and the remaining 20% for testing.This separation ensures that the model can be trained on a subset of the data and evaluated on an independent subset, providing a realistic assessment of its performance.

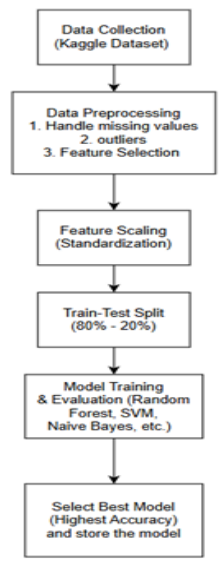

Fig 2: Flow Diagram of Proposed Method (Phase 1)

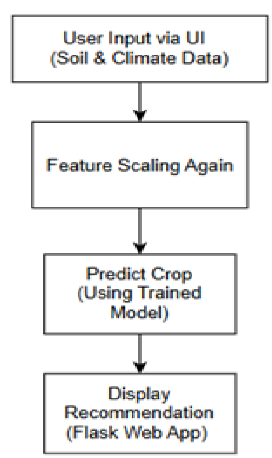

Fig 3: Flow Diagram of Proposed Method (Web Based Interface Phase 2)

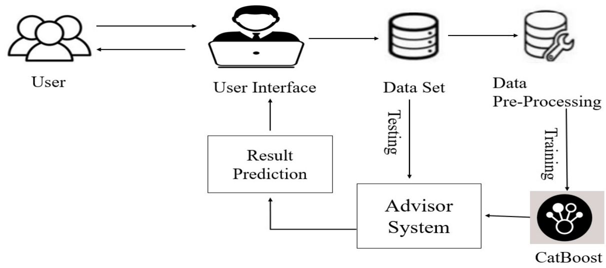

Fig 4: Proposed Architecture

V. Feature Scaling

Feature scaling is an essential preprocessing step that normalizes the range of feature values, ensuring that all features contribute equally to the model's learning process. In the "Smart Crop Advisor" system, Standard Scalar Normalization is employed to rescale the features with a mean of 0 and a standard deviation of 1. This normalization helps to address issues related to features with different units or scales, which can otherwise lead to biased model performance. For instance, features with larger numerical ranges can dominate the learning process, overshadowing features with smaller ranges. By applying Standard Scalar Normalization, each feature is standardized to have a consistent distribution around zero, thus enhancing the stability and effectiveness of the machine learning algorithms. This preprocessing step is crucial for achieving accurate and reliable predictions from the models, as it ensures that all input features are treated uniformly during the training and evaluation phases.

VI. Model Training and Evaluation

Once the data has been preprocessed and scaled, the next step is to train and evaluate various ML models to determine which offers the best crop recommendations based on environmental conditions. In this research, nine models are used: Gradient Boosting Machines (GBM), CatBoost, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Nearest Centroid (NC), Decision Tree, Support Vector Machine (SVM), Naive Bayes and Random Forest. Each model contributes uniquely to the recommendation process by leveraging different approaches. GBM is an ensemble technique that builds decision trees sequentially, where each tree corrects the errors of the previous ones, leading to high predictive accuracy and robustness. CatBoost is another powerful gradient-boosting algorithm that efficiently handles categorical features, providing strong performance by reducing overfitting and improving accuracy. Both GBM and CatBoost excel in modeling complex, nonlinear relationships between environmental features like soil type, rainfall, and temperature.

LDA is used to find linear combinations of features that best separate different crop classes, reducing dimensionality and projecting data into a lower-dimensional space. QDA extends LDA by assuming that each class follows a different covariance structure, allowing for more flexible decision boundaries and better classification performance when crop classes exhibit nonlinear separability. Nearest Centroid (NC) offers a simple yet effective approach by classifying crops based on proximity to the average environmental conditions for each crop class, making it computationally efficient and useful for large-scale or real-time applications.

The Decision Tree model creates a series of splits based on environmental conditions, offering easy-to-interpret decisions and performing well with low-variance data. The Support Vector Machine (SVM) finds an optimal hyperplane that separates different crop classes, excelling in high-dimensional spaces and ensuring robust performance with nonlinear relationships. Naive Bayes, a probabilistic model, uses Bayes' theorem with the assumption of feature independence, making it efficient and effective, particularly with large datasets and when the features are loosely correlated.Lastly, Random Forest, an ensemble of decision trees, enhances prediction stability and accuracy by averaging the results from multiple trees, reducing overfitting. Together, these models offer diverse methodologies that account for the variability in environmental factors affecting crop growth, leading to accurate, reliable, and efficient crop recommendations.

VII. Model Selection

After training and evaluating all the machine learning models, selecting the most suitable model is based on performance metrics such as accuracy and confusion matrix. Accuracy measures the proportion of correct predictions made by the model, while the confusion matrix provides detailed information on the types of errors true positives, true negatives, false positives, and false negatives. Among the models tested, the one with the highest accuracy and the most favorable confusion matrix results is selected. This chosen model is considered the best fit for making accurate crop predictions based on the given environmental conditions. The selection process ensures that the final model is reliable and performs well in providing crop recommendations, which is crucial for the effectiveness of the "Smart Crop Recommendation."

A. Recommendation System

The Recommendation System utilizes the selected ML model to provide crop suggestions based on user inputs from a web interface. The system takes various environmental parameters—Nitrogen (N), Phosphorus (P), Potassium (K), temperature, humidity, pH, and rainfall—from the user. It normalizes these inputs using the same scaling method applied during preprocessing.

The input data from the web interface is normalized and then fed into the trained machine-learning model by using Flask to generate a prediction. This prediction, represented by a numerical value, is mapped to the corresponding crop label using a predefined dictionary. The system's output is a specific crop recommendation tailored to the user's environmental conditions displayed on the web interface which is shown in fig 4 and fig 5. This approach ensures that the ”Smart Crop Advisor” delivers accurate and relevant crop suggestions, aiding users in making informed agricultural decisions.

Fig 5 : Recommendation 1 (Rice)

Fig 6 : Recommendation 2 (Coffee)

B. Dataset Description

The ”Crop recommendation” dataset available in Kaggle is used which includes various features such as Nitrogen (N), Phosphorus (P), Potassium (K), temperature, humidity, pH, and rainfall, which are essential for understanding the conditions that influence crop suitability. Additionally, the dataset contains crop labels that indicate which crops are most suitable under specific conditions. This dataset provides a comprehensive basis for training and evaluating machine learning models, enabling the system to learn patterns and relationships between environmental factors and crop types. The quality and representativeness of the data are vital as they directly impact the accuracy and reliability of the predictions made by the ”Smart Crop Recommendation.”

VIII. Experimental Setup

Different The experimental setup for the Smart Crop Advisor involved training multiple machine learning models using the Kaggle Crop Recommendation dataset, which contains soil nutrients (N, P, K), climate factors (temperature, humidity, rainfall), and soil pH to predict the most suitable crop. Data preprocessing included outlier removal using Interquartile Range (IQR), feature-target separation, and an 80-20 train-test split. Feature scaling was applied using Standard Scalar Normalization to ensure uniform data distribution. Nine machine learning models Random Forest, Naive Bayes, CatBoost, GBM, QDA, SVM, Decision Tree, LDA, and Nearest Centroid were trained and evaluated using cross-validation and hyperparameter tuning. Model performance was assessed using accuracy, precision, recall, F1-score, and confusion matrices, with Random Forest achieving the highest accuracy of 99.7%. Finally, the best-performing model was integrated into a Flask-based web application, allowing users to input environmental parameters and receive real-time crop recommendations.

IX. RESULTS AND DISCUSSION

Table 1 presents the performance metrics of various machine learning models trained on the Crop Recommendation Dataset available on Kaggle. These models were evaluated to recommend crops efficiently based on soil and weather conditions, with metrics including accuracy, precision, recall, and F1 score. Among the models, Random Forest achieved the highest accuracy of 99.7%, with nearly perfect precision, recall, and F1 score, indicating its robustness in predicting the most suitable crops. Naive Bayes and CatBoost also performed exceptionally well, both achieving high validation accuracies of 99.2%. On the other hand, the Nearest Centroid (NC) model had the lowest accuracy at 89.8%, though it still demonstrated a reasonable performance. Models like GBM and QDA also showed strong results with accuracies of 98.9%. These results highlight the varying capabilities of different models in predicting crop recommendations based on the dataset, providing valuable insights for future research and applications in agriculture.

Table 1. Crop Recommendation Results

S. No.	Model	Accuracy	Precision	Recall	F1 Score
1	Random Forest	99.7%	0.997	0.997	0.997
2	Naive Bayes	99.2%	0.992	0.992	0.992
3	CatBoost	99.2%	0.991	0.992	0.991
4	GBM	98.9%	0.990	0.989	0.989
5	QDA	98.9%	0.989	0.989	0.989
6	SVM	97.7%	0.982	0.977	0.974
7	Decision Tree	97.2%	0.974	0.972	0.970
8	LDA	97.2%	0.972	0.972	0.972
9	NC	89.8%	0.911	0.898	0.898

Confusion Matrices of various trained models are shown They show the classification capacity and efficiency of each model in learning the dataset and predicting.Table 1 presents the performance metrics of various ma- chine learning models trained on the Crop Recommendation Dataset available on Kaggle. These models were evaluated to recommend crops efficiently based on soil and weather conditions, with metrics including accuracy, precision, recall, and F1 score.Among the models, Random Forest achieved the highest accuracy of 99.7%, with nearly perfect precision, recall, and F1 score, indicating its robustness in predicting the most suitable crops. Confusion Matrices of various trained models are shown in Fig7-Fig8. They show the classification capacity and efficiency of each model in learning the dataset and predicting.

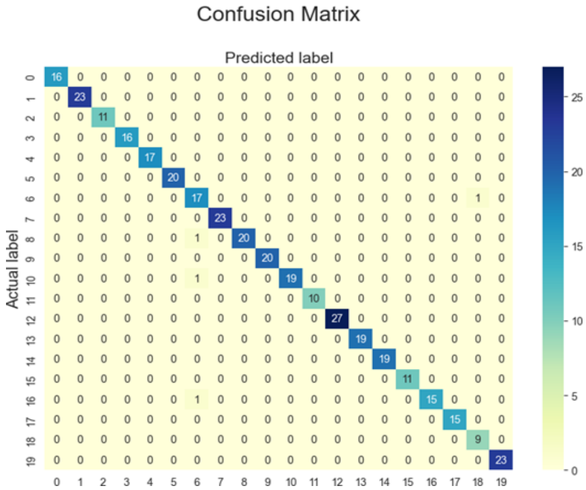

Fig 7 : Random Forest

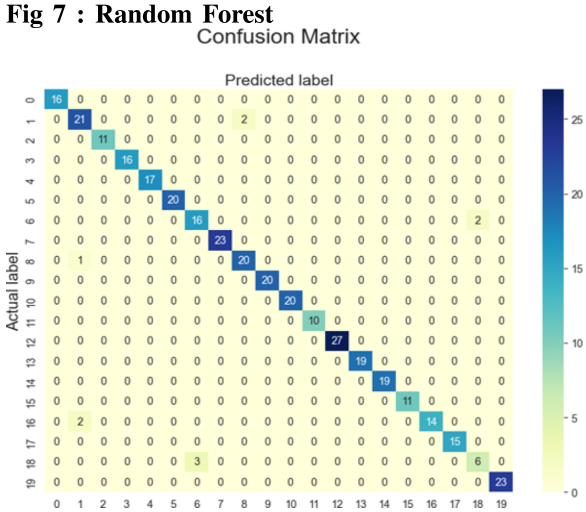

Fig 8 : Catboost

The below bar chart demonstrates the performance of accuracy for different machine learning models when run on a standard dataset. Compared models are GBM, CatBoost, LDA, QDA, NC, Decision Tree, SVM, Naïve Bayes, and Random Forest.

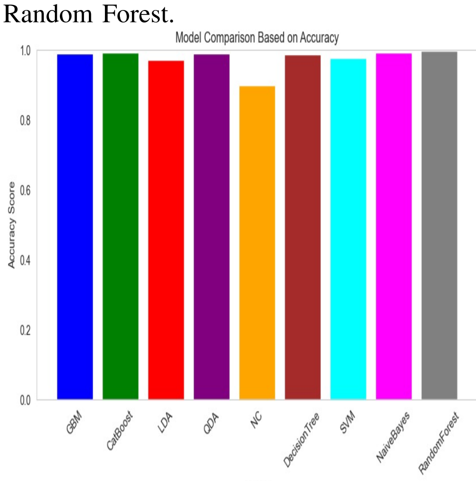

Fig 10 : Model Comparison Based on Accuracy.

X. Conclusion

The "Smart Crop Advisor" demonstrates the potential of machine learning in agricultural decision-making. The proposed system helps farmers make informed decisions, thereby improving crop yields, reducing resource usage, and supporting sustainable agriculture. Future work may focus on integrating IoT sensors for real-time data input and expanding the system to include fertilizer recommendations.

XI. Future work

Future work for the Smart Crop Advisor can focus on enhancing its predictive capabilities and expanding its real-world applicability. Integrating real-time data collection using IoT sensors and remote sensing technology would improve accuracy by continuously updating soil and weather parameters. Additionally, incorporating deep learning models such as Convolutional Neural Networks (CNN) could enhance spatial data analysis, enabling the system to process satellite images, drone-captured field images, and soil scans for better crop and soil health assessment. Expanding the dataset to include market trends, pest risks, and economic factors could provide farmers with more profitable and sustainable crop recommendations.

Developing a mobile application alongside the web-based platform would increase accessibility for farmers, especially in rural areas.

XII. References

[1] Pawar, Shubham, et al. "Smart farming using machine learning." *Smart Comput* (2021): 73-81.

[2] Rao, Madhuri Shripathi, et al. "Crop prediction using machine learning." *Journal of Physics: Conference Series*. Vol. 2161. No. 1. IOP Publishing, 2022.

[3] Prity, F.S., Hasan, M.M., Saif, S.H. et al. "Enhancing Agricultural Productivity: A Machine Learning Approach to Crop Recommendations." *Hum-Cent Intell Syst* (2024).

[4] Sharma, Priyanka, Pankaj Dadheech, and AV Senthil Kumar Senthil. "AI-Enabled Crop Recommendation System Based on Soil and Weather Patterns." *Artificial Intelligence Tools and Technologies for Smart Farming and Agriculture Practices*. IGI Global, 2023. pp. 184–199.

[5] Dahiphale, Devendra, et al. "Smart farming: Crop recommendation using machine learning with challenges and future ideas." *Authorea Preprints* (2023).

[6] Garg, Disha, and Mansaf Alam. "An effective crop recommendation method using machine learning techniques." *International Journal of Advanced Technology and Engineering Exploration*, vol. 10, no. 102, 2023, p. 498.

[7] Ghosh, Asit, et al. "A Comprehensive Crop Recommendation System Integrating Machine Learning and

Deep Learning Models." *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)*. IEEE, 2024.

[8] ReenaSri, S., U. Dhanushree, and K. Sivadhanu. "Crop Recommendation Systems using Machine Learning Algorithms." *2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)*. IEEE, 2023.

[9] Parameswari, Ponnusamy, N. Rajathi, and K. J. Harshanaa. "Machine learning approaches for crop recommendation." *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*. IEEE, 2021.

[10] Mewar, Amritya, et al. "Design of Web Based Recommendation System for Farmers using Machine Learning." *2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET)*. IEEE, 2024.

[11] Iniyan, S., et al. "Crop and Fertilizer Recommendation System Applying Machine Learning Classifiers." *2023 International Conference on Emerging Research in Computational Science (ICERCS)*. IEEE, 2023.

[12] Shedthi, Shabari, et al. "Machine Learning Techniques in Crop Recommendation based on Soil and Crop Yield Prediction System–Review." *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE)*. IEEE, 2022.

[13] Islam, Tanhim, Chisty, Tanjir Alam, and Chakrabarty, Amitabha. "A Deep Neural Network Approach for Crop Selection and Yield Prediction in Bangladesh." *arXiv preprint*, August 2021.

[14] Gupta, Shloka, Chopade, Akshay, Jain, Nishit, and Bhonde, Aparna. "Farmer's Assistant: A Machine Learning-Based Application for Agricultural Solutions." *arXiv preprint*, April 2022.

[15] Shahhosseini, Mohsen, Hu, Guiping, Archontoulis, Sotirios V., and Huber, Isaiah. "Coupling Machine Learning and Crop Modeling Improves Crop Yield Prediction in the US Corn Belt." *arXiv preprint*, July 2020.

[16] Suresh, A., Kumar, P. Ganesh, and Ramalatha, M. "Prediction of Major Crop Yields of Tamilnadu Using K-Means and Modified KNN." *Proceedings of the 2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, October 2018.

[17] Khaki, Saeed, and Wang, Lizhi. "Crop Yield Prediction Using Deep Neural Networks." *arXiv preprint*, February 2019.

[18] Sainte Fare Garnot, Vivien, Landrieu, Loic, Giordano, Sebastien, and Chehata, Nesrine. "Time-Space Tradeoff in Deep Learning Models for Crop Classification on Satellite Multi-Spectral Image Time Series." *arXiv preprint*, January 2019.

[19] Paul, Monali, Vishwakarma, Santosh Kumar, and Verma, Ashok. "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach." *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, December 2015.

[20] Ahamed, A.T.M.S., et al. "Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh." *2015 IEEE/ACIS*.

[21] Pudumalar, S., et al. "Crop Recommendation System for Precision Agriculture." *2016 Eighth International Conference on Advanced Computing (ICoAC)*, January 2017.

[22] Rajak, R.K., et al. "Crop Recommendation System to Maximize Crop Yield Using Machine Learning Technique." *International Journal of Research in Advent Technology*, 2017.

[23] Sanghvi, Yash, et al. "Comparison of Self Organizing Maps and Sammon's Mapping on Agricultural Datasets for Precision Agriculture." *2015 International Conference on Computing Communication Control and Automation (ICCUBEA)*, August 2015.

[24] Mehta, Viraj, et al. "A Machine Learning Approach to Foretell the Probability of a Crop Contracting a Disease." *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, August 2018.