Grenze International Journal of Engineering and Technology, June Issue

Intelligent Farming: Crop Recommendation System Powered by Decision Tree Classification

Sridhar Chintala[1], Ashok Nimmala[2], Vishwanath Bijalwan[3] and Deep Shekhar Acharya[4]
[1]School of CS&AI, SR University, Warangal, India
Email: dr.sridharchintala@gmail.com, ashoknimmala050@gmail.com, vishwanath.bijalwan@gmail.com
[2-4]Dept. of EEE, Birla Institute of Technology, Mesra, Ranchi, India
Email: dsacharya@bitmesra.ac.in

***Abstract*—Farm technologies such as crop guidance tools remains very useful in enhancing and improving the yields in farming. The objective of this research is therefore to design a decision tree crop advice mechanism to aid farmers identify high return crops that will thrive in the current land and climate of a given region. This method employs active use of new data analysis to increase farm yields, and reduce on the impact farms have on the environment. It is originate from a website, nitrogen, phosphorus potassium temperature and humidity, ph and rainfall etc are shown there. Decision tree is constructed for analysing the data and to identify patterns, relations and clusters. The suggested method was compared against the existing algorithms like K-Means, K-Medoid; Support Vector Machine (SVM). Decision Tree algorithm had the highest accuracy, at 99%.**

***Index Terms*—Crop Recommendation Systems, Soil Characteristics, Trial-and-Error Farming, Climate Variability.**

I. INTRODUCTION

Agriculture as a sector, plays a significant role in the country's economy, culture and food security as it employs a good proportion of its population [1]. India is agricultural based contributing a lot to GDP through income generating crops such as paddy, grains, maize which are used for importation and local consumption. Crop and Livestock DIVERSIFICATION [2], in Agriculture is crucial to REAFFIRM and SUPPORT rural development, and striving towards sustainable living in the rural lands. Moreover, Indian agriculture's position is well prevalent through exports of rice, ginger, spices and cotton etc., which adds to the development of the country and its economy. As result of this, crop recommendation systems have proved vital in improving agricultural productivity. These systems allow the farmers to grow crops appropriate for given zones depending on features such as soil acidity, climate, moisture and demand. With widespread types of crops grown all over India and varied climate conditions these systems help in proper distribution of crops and hence, efficient use of resources leading to better productivity in the field of agriculture. Current crop advisory services use highly developed techniques including but not limited to Artificial Intelligence (AI), Machine Learning (ML)[3], satellite imagery for data analysis. This not only looks at saving on costs such as fertilizer and pesticide use or irrigation but also considers ways to minimize harm to the environment given the recommendations on the use of these resources are now made to individual farmers based on real-time use.

Grenze ID: 01.GIJET.11.2.232_20

From the discussed research, the application of applicable ML algorithms [4], including the Random Tree, CHAID, K-Nearest Neighbour and the Naive Bayes has enhanced the reliability of crops recommendation. Majority voting based methodologies incorporating ensemble methods have shown improved accuracy and other performance measures to transform agriculture into a more fruitful and sustainable environment. The literature survey in [5] explores various methods for crop disease detection, focusing on plants like cotton and grapes. It highlights a Decision Tree Classifier for detecting potential cotton crop diseases using factors such as temperature and soil moisture. The study emphasizes improved farming performance and features a user-friendly Android application that provides real-time results to farmers.In [6], a systematic approach to crop recommendation using machine learning (ML) is detailed. Relevant datasets were collected from Indian districts, cleaned, and balanced using the SMOTE technique to address class imbalances. Thirteen classifiers were evaluated on metrics for example accuracy, precision, and recall, with the Stochastic Gradient Descent Classifier (SGDC) achieving 100% accuracy after applying SMOTE. An ontology-based agricultural advisory system, PCT-O (Pests in Crops and their Treatments Ontology), is proposed in [7]. It addresses challenges in managing pest-related information, often scattered and heterogeneous. PCT-O integrates data from taxonomies, chemical databases, and agricultural guides, semantically linking crop-pest-treatment relationships. It performs tasks such as pest identification, treatment recommendations, and policy comparisons by region. Despite challenges with data heterogeneity and alignment, PCT-O demonstrates potential for enhancing decision-making in crop protection.

The study [8] surveys existing ML-based crop recommendation systems, addressing challenges like poor crop selection, lack of soil awareness, and climate variability in Indian agriculture. It reviews models utilizing techniques such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forests (RF), and ensemble methods like Majority Voting, which excel in crop yield estimation. The survey emphasizes incorporating environmental, soil, and economic factors to enhance system effectiveness. Future recommendations include integrating geospatial analysis and refining algorithms for precision. This study implemented a crop recommendation system using SVM, Naive Bayes, and ANN with Majority Voting, improving prediction accuracy.

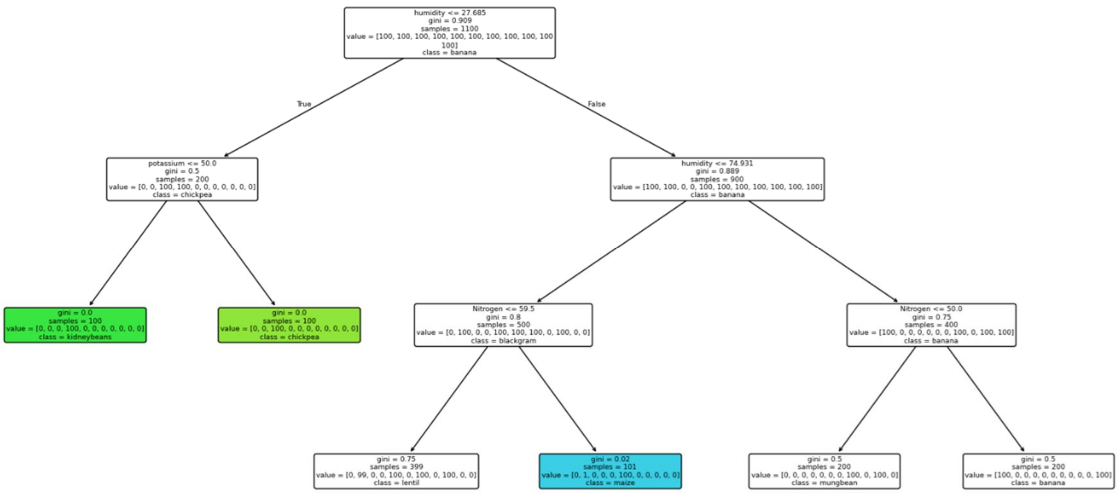

Figure 1. Decision tree for crop recommendation dataset

II. WORKING MODEL

The crop recommendation system[9], uses a variety of agricultural and environmental factors to recommend the most suitable crops for a given piece of land. In this model, we employ ML techniques such as DecisionTree Classification, as it is well-suited for interpreting the relationships between input features like soil quality, climate and crop choices, respectively. The dataset is collected from [10], which contains information on soil nutrient status, environmental factors (temperature, humidity and rainfall) and soil pH. It also includes crop labels indicating the suitable crops for the respective features. The dataset is preprocessed before extracting the features. The normalization is applied all the numeric values of features of the data. Finally label encoding is performed on the categorical features. The data split of 80% for training , and 20% test, respectively .The decision tree is implemented upon data for classification.

The Decision tree classifier [11] is supervised machine learning algorithm that changes the data into tree –like structure. It operates on four points, which are: Feature Selection where, at each node finds out the best feature to split the data into similar clusters. The conditions for the formation of these splits is usually based on Entropy[12]. Later, recursive Splitting is used to subdivide the dataset recursively into subsets (forming a tree) until the stopping metric is reached. After a tree has been trained, new data points traverse the tree according to their features until reaching a leaf node. Leaf nodes represents a predicted crop type.

A. Feature Importance Plot

It shows the contribution of each attribute to the prediction. For a feature ***f***, the importance I(f) is defined as:

$$I(f) = \sum_{t \in T(f)} \frac{N_t}{N} \left(Impurity_{Parent} - WeightedAverageImpurity_{Child} \right)$$

T(f) is the Nodes where the feature ***f*** is used, Nt: Number of samples at node *t* and N: Total number of samples in the dataset.

II. Results and Discussion

Figure 1 visualizes the Decision Tree model, highlighting how decisions are made based on features like Nitrogen, Phosphorus, Potassium, Temperature, Humidity, and Soil pH. Internal nodes represent questions about feature values, with branches providing answers, such as "Is Nitrogen > 30?" or "Is Temperature < 25°C?" Figure 2 displays a confusion matrix, a key tool for evaluating classification models. It compares the predicted crop classes against the actual ones from the test data, with rows representing actual classes and columns showing predicted ones. Figure 3 presents a feature importance graph, illustrating the contribution of each feature to the model's decisions. Nutrients like Nitrogen, Phosphorus, and Potassium play a critical role in determining suitable crops, as they are vital for growth. Feature importance is calculated by the improvement in impurity during the tree-building process, emphasizing how key parameters influence the model's output.

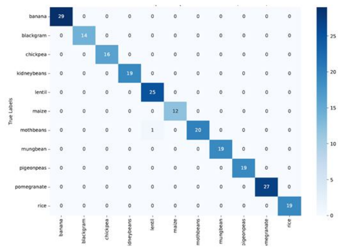

Figure 2. Confusion matrix of proposed decision tree

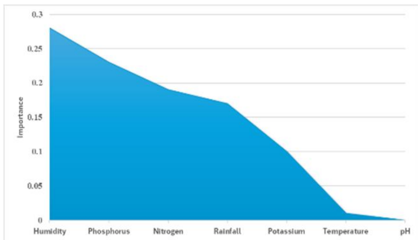

Figure 3.Feature importance plot of decision tree

Figure 4 presents the precision-recall curve, emphasizing the balance between precision (fewer incorrect predictions) and recall (broader crop suggestions). Figure 5 displays the ROC curve, contrasting the true positive and false positive rates to evaluate the classifier's performance. A high AUC reflects the model's strong execution. The model demonstrated excellent accuracy, along with a well-balanced precision and recall, supported by a solid F1 score in Table I, confirming its practicality for real-world use. However, predicting crops with similar environmental factors proved challenging. Although the system provides valuable insights, its success relies on the quality, accuracy, and timely availability of input data, such as soil tests and environmental measurements.

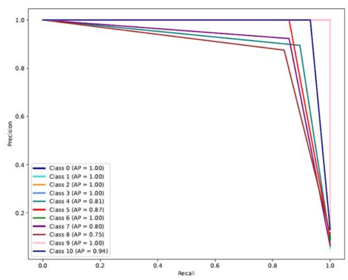

Figure 4. Precision recall curve

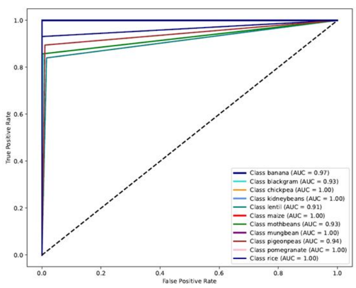

Figure 5. ROC curve

TABLE I. PERFORMANCE METRICS FOR DIFFERENT MODELS

Metric	SVM	K-Means	K-Medoid	Decision Tree
Accuracy	0.85	0.85	0.77	0.99
Precision	0.91	0.91	0.86	0.99
Recall	0.85	0.85	0.77	0.99
F1 Score	0.82	0.82	0.71	0.99

IV. CONCLUSION

The crop recommendation system has remarkable potential in supporting precision agriculture. Farmers and agronomists can, however, make use of the methods successfully conducted using decision trees, feature importance plots, and evaluation curves. Although the modeling presented in this study is satisfactory, more efforts in the near future will make the model even more flexible and robust increasing its usefulness for sustainable farming practices.

REFERENCES

[1] Mathur, A. S., Das, S., & Sircar, S. (2006). Status of agriculture in India: trends and prospects. Economic and political weekly, 5327-5336.

[2] P Parthasarathy Rao and AJ Hall. Importance of crop residues in crop–livestock systems in india and farmers'perceptions of fodder quality in coarse cereals. Field Crops Research, 84(1-2):189–198, 2003.

[3] Zhou, Z. H. (2021). Machine learning. Springer nature.

[4] Aized Amin Soofi and Arshad Awan. Classification techniques in machine learning: applications and issues. Journal of Basic & Applied Sciences, 13:459–465, 2017.

[5] Chopda, J., Raveshiya, H., Nakum, S., & Nakrani, V. (2018, January). Cotton crop disease detection using decision tree classifier. In 2018 International Conference on Smart City and Emerging Technology (ICSCET) (pp. 1-5). IEEE.

[6] Senapaty, M. K., Ray, A., & Padhy, N. (2024). A decision support system for crop recommendation using machine learning classification algorithms. Agriculture, 14(8), 1256.

[7] avier Lacasta, F Javier Lopez-Pellicer, Borja Espejo- Garc´ıa,JavierNogueras-Iso,andFJavierZarazaga-Soria. Agriculturalrecommendationsystemforcropprotection. Computers and Electronics in Agriculture, 152:82–89, 2018.

[8] Jadhav, R., & Bhaladhare, P. (2022). A machine learning based crop recommendation system: a survey. Journal of Algebraic Statistics, 13(1), 426-430.

[9] Pradeepa Bandara, Thilini Weerasooriya, T Ruchirawya,W Nanayakkara, M Dimantha, and M Pabasara. Crop recommendation system. International Journal of Computer Applications, 975:8887, 2020.

[10] Nitesh,H.Crop recommendation dataset,2020.

[11] Kaushik K , Rana et al.A survey on decision tree algorithm for classification.International journal of Engineering development and research, 2(1):1–5, 2014.

[12] Berthold Bein. Entropy. Best Practice & Research Clinical Anaesthesiology, 20(1):101–109, 2006