International Journal of
Human Computations & Intelligence
Vol. 03, Issue. 02, Mar/Apr 2024
RESEARCH ARTICLE OPEN ACCESS

An efficient crop recommendation system using machine learning mechanisms

Kiran Kumar P N[1] . Bhavya R A[1] . Dhanushree A N[1] . Nagarjuna G R[2]

[1]Dept. of Computer Science and Engineering,
SJC Institute of Technology, VTU, Chikkaballapur, India

[2]Dept. of Artificial Intelligence and Machine Learning,
SJC Institute of Technology, VTU, Chikkaballapur, India

DOI: **10.5281/zenodo.11441063**
Received: 29 January 2024 / Revised: 18 February 2024 / Accepted: 13 April 2024
©Milestone Research Publications, Part of CLOCKSS archiving

Abstract — In India, the job situation and economy are significantly influenced by agriculture. However, a common problem for Indian farmers is choosing the wrong crops for their land, which lowers yield. In addition to having an impact on individual farmers' earnings, this problem has larger ramifications for the country's food security and is a factor in the surge in farmer suicides. Proactive steps are needed to address these issues, such as recommending appropriate crops based on soil tests before sowing. A crop recommendation system that incorporates machine learning algorithms is one suggested way to address this. The objective is to reduce farmer losses and increase total productivity by evaluating the profitability of individual crops. Various soil factors are used to identify the most suited crops through the use of machine learning algorithms for classification. To verify dependable results, the efficacy of this method is tested by calculating accuracy and confusion matrix metrics. The objective is to equip farmers with the knowledge necessary to make wise decisions by strategically applying cutting-edge algorithms and data analysis. This will ultimately promote sustainable agricultural practices and solve the sector's problems.

Index Terms – Agriculture, machine learning algorithms, Confusion Matrix.

I. INTRODUCTION

India's economy and jobs are based primarily on agriculture, yet while being one of the world's top producers, India struggles with low productivity. The agricultural landscape has changed due to globalization, necessitating the development of creative solutions to maintain livelihoods and increase output. Inappropriate crop selection is a major issue that reduces productivity [1] [2]. To address this, the proposed effort is to create a recommendation system that would direct precise crop selection, increasing production and decreasing incorrect decisions. This project includes crop prediction,

MILESTONE RESEARCH.IN OPEN ACCESS

ISSN (Online): 2583-5696
Int. Jr. of Hum Comp. & Int.

classification, model training, and evaluation, with an emphasis on utilizing cutting-edge technologies to support agricultural practices. Although conventional agricultural practices continue to be used, adopting data-driven insights could maximize crop production in the face of changing weather and soil conditions. By extending the body of knowledge on crop recommendation algorithms, this effort aims to equip Indian farmers with the knowledge and skills necessary to make well-informed decisions, thereby bringing in a new era of sustainable and efficient agriculture [3][4].

It is essential to integrate contemporary technologies with conventional farming methods to maximize agricultural productivity and satisfy expanding consumer expectations. The usage of machine learning algorithms has given solutions to various human applications [5] for classifications and predictions [6]. Developing an efficient crop recommendation model requires the application of state-of-the-art machine learning methods. Authorities can maximize projected yields by making well-informed judgments about crop cultivation, fertilization, and seed selection thanks to artificial intelligence and machine learning. To improve yield estimates in agriculture and achieve higher precision, this research uses an ensemble machine learning technique. Diverse cutting-edge tactics are used to increase output, all directed toward satisfying necessities [7]. To increase agricultural productivity, advice for premium seeds, fertilizers, and implements are crucial. To prevent negative effects, farmers must choose the appropriate crops and production techniques, highlighting the significance of their decisions in maximizing raw materials and satisfying nutritional needs [8]. The remainder of the paper is structured in the same way as section 2, which talks about the other relevant research that has been done on the subject. Section 3 elaborates on the proposed model with all the steps involved in it. Section 4 focuses on the results produced from the experimentation and the related discussions. Finally, the conclusion of the paper is presented in section 5.

II. RELATED WORKS

A framework in the study [9] combines model-based techniques like the KNN algorithm, Support Vector Machine (SVM), and Brute method with collaborative filtering. According to their results, KNN performed better in terms of prediction accuracy than SVM and the Brute algorithm. To improve forecast accuracy even more, an ensemble technique with majority voting was the topic of Paper [10]. SVM, AdaBoost, and artificial neural networks (ANN) were all used in this ensemble approach. High accuracy values were obtained by using KNN and Naive Bayes as base learners and the majority voting approach. Real-time soil condition measurements were obtained through IoT deployment, which guided crucial activities for higher production. Crop suitability and fertilizer recommendations were reported in the paper [11], which used Random Forest (RF), KNN, and SVM to achieve accuracy rates of 89% and 80%, respectively. To increase model accuracy, an ensemble of machine algorithms was merged [12].

Farmers were able to reap the benefits of a modified technique that used NB classification to recommend crops and fertilizers. Gradient Descent was utilized in Paper [13] to train the prediction models and increase forecast accuracy. The models were created using RF and NB. SVM, NB, MLP, J48(C4.5), and JRip were among the classifiers used for soil analysis. To choose the best plants based on soil and environmental conditions, the paper [14] addressed the feature selection and fragmentation of essential components. By removing irrelevant data and reducing the amount of features without sacrificing important information, feature selection techniques improve prediction accuracy. This study highlights how machine learning techniques might boost agricultural yield, providing a viable path for improving crop selection and productivity.

A recommender system was presented in the paper [15] to choose the best crop type based on a number of variables, such as irrigation availability, fertilizer type, and weather forecast. They found that the best crop type is dependent on both the overall amount of rainfall and the existing training model. A deep learning-based technique called PSO-Modified DNN was presented in a different study [16] and shown to be more accurate in predicting agricultural production than a number of machine learning methods. In order to provide extremely accurate crop yield suggestions, this method uses an ensemble model using a majority vote strategy that combines SVM with ANN, Random Tree, and NB-classifier. The study [17] estimated crop yield using fuzzy logic and cosine similarity, and advised seeds, pesticides, and equipment based on farming practices and geographic location. Similarly, [18] suggested a precision agriculture strategy for the best crop selection through the use approaches. Using map-reduce with KNN, the paper [19] effectively recommended crops by taking into account geographical variables such soil texture, hue, drainage, and climate conditions. Concurrently, [20] proposed an ensemble model for precision agriculture by utilizing diverse machine learning methods, while created a fuzzy expert multi-layer crop recommendation system. Using a parallel fuzzy rule base system, the paper [21] concentrated on important factors including geography and soil texture data impacting continuous crop production. Finally, [22] mined agricultural data for crop production enhancement factors by using PAM, CLARA, DBSCAN, and Multiple Linear Regression and then used the KNN algorithm to recommend crop types based on efficacy criteria.

III. PROPOSED METHOD

The flow of the proposed work is shown in Figure 1. The proposed work is elaborated by considering five major machine-learning algorithms such as decision trees, naïve bayes, logistic regression, k-nearest neighbour and random forest the recommendation of crops.

Dataset Description

The collection includes 22 crop varieties. For every crop, it includes 100 data points that represent various weather and soil consistency situations. The dataset, which was assembled from Kaggle, facilitates the optimization of agricultural production by including factors like N, P, K, temperature, humidity, pH, rainfall, and crop variety. However, because it is a subset, there could not be as many accurate predictions. The entire dataset was shuffled to equally distribute the successive crop kinds to guarantee homogeneity. Crop variations are divided equally across the training and test sets so that every feature is represented fairly for thorough analysis.

Preprocessing Phase

Data is preprocessed at this step and then supplied into the machine learning algorithm. Row IDs and other columns deemed unnecessary for prediction are eliminated. The majority of machine learning algorithms are limited to working with numerical data, hence non-numeric items (like labels) in the dataset must be translated. We call this conversion process label encoding. By giving each class a distinct numerical value using label encoding, the algorithm is able to process the data efficiently. The fit_transform function of the pandas library is used to transform the columns in our dataset—whose labels are initially non-numeric—into their matching encoded values. The dataset is then divided for training and testing usage by the machine learning models.

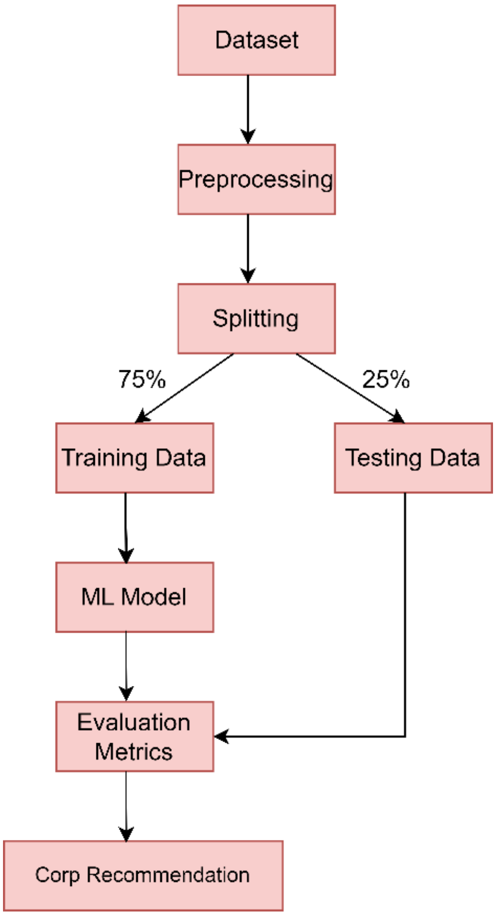

FIG. 1: PROPOSED APPROACH

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed work has been implemented in Google Colab and analyzed the performance of considered machine learning algorithms. The performance is measured by considering the accuracy, precision, recall and F1 score of each classification algorithm. Figures 2 to 6 show the classification reports having the precision, recall, f1 score and accuracy of each machine algorithm for the considered type of crop data. Table 1 shows the summary of the performance parameters of each machine-learning algorithm. And the comparison of accuracy is highlighted in Figure 7. For a machine learning algorithm, Random Forest frequently shows better accuracy than many other algorithms. By combining predictions from several decision trees, its ensemble learning method helps to reduce overfitting and enhance generalization. This makes it extremely useful for tasks involving regression and classification, particularly when working with complicated interactions or high-dimensional data. Nevertheless, several variables, including the particulars of the dataset, the available computing power, and the nature of the issue at hand, ultimately determine which approach is best suited. Therefore, even though Random Forest usually produces results with high accuracy, careful evaluation of other algorithms is still necessary to choose the best model.

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	24
banana	1.00	1.00	1.00	19
blackgram	0.61	1.00	0.75	20
chickpea	1.00	1.00	1.00	26
coconut	0.93	1.00	0.96	27
coffee	1.00	1.00	1.00	26
cotton	1.00	1.00	1.00	24
grapes	1.00	1.00	1.00	23
jute	0.50	0.03	0.06	31
kidneybeans	0.00	0.00	0.00	25
lentil	0.62	1.00	0.76	26
maize	1.00	1.00	1.00	26
mango	1.00	1.00	1.00	31
mothbeans	0.00	0.00	0.00	25
mungbean	1.00	1.00	1.00	29
muskmelon	1.00	1.00	1.00	26
orange	1.00	1.00	1.00	34
papaya	1.00	1.00	1.00	23
pigeonpeas	0.53	1.00	0.70	24
pomegranate	1.00	1.00	1.00	19
rice	0.46	0.96	0.62	25
watermelon	1.00	1.00	1.00	17
accuracy			0.85	550
macro avg	0.80	0.86	0.81	550
weighted avg	0.80	0.85	0.80	550

FIG. 2: CLASSIFICATION REPORT OF DECISION TREES ALGORITHM

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	24
banana	1.00	1.00	1.00	19
blackgram	1.00	1.00	1.00	20
chickpea	1.00	1.00	1.00	26
coconut	1.00	1.00	1.00	27
coffee	1.00	1.00	1.00	26
cotton	0.96	1.00	0.98	24
grapes	1.00	1.00	1.00	23
jute	0.89	1.00	0.94	31
kidneybeans	1.00	1.00	1.00	25
lentil	1.00	1.00	1.00	26
maize	1.00	0.96	0.98	26
mango	1.00	1.00	1.00	31
mothbeans	1.00	1.00	1.00	25
mungbean	1.00	1.00	1.00	29
muskmelon	1.00	1.00	1.00	26
orange	1.00	1.00	1.00	34
papaya	1.00	1.00	1.00	23
pigeonpeas	1.00	1.00	1.00	24
pomegranate	1.00	1.00	1.00	19
rice	1.00	0.84	0.91	25
watermelon	1.00	1.00	1.00	17
accuracy			0.99	550
macro avg	0.99	0.99	0.99	550
weighted avg	0.99	0.99	0.99	550

FIG. 3: CLASSIFICATION REPORT OF NAÏVE BAYES ALGORITHM

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	24
banana	1.00	1.00	1.00	19
blackgram	0.88	0.70	0.78	20
chickpea	1.00	1.00	1.00	26
coconut	1.00	1.00	1.00	27
coffee	1.00	1.00	1.00	26
cotton	0.88	0.92	0.90	24
grapes	1.00	1.00	1.00	23
jute	0.85	0.94	0.89	31
kidneybeans	1.00	1.00	1.00	25
lentil	0.87	1.00	0.93	26
maize	0.92	0.88	0.90	26
mango	0.97	1.00	0.98	31
mothbeans	0.81	0.84	0.82	25
mungbean	1.00	1.00	1.00	29
muskmelon	1.00	1.00	1.00	26
orange	1.00	1.00	1.00	34
papaya	1.00	0.91	0.95	23
pigeonpeas	1.00	1.00	1.00	24
pomegranate	1.00	1.00	1.00	19
rice	0.91	0.80	0.85	25
watermelon	1.00	1.00	1.00	17
accuracy			0.96	550
macro avg	0.96	0.95	0.96	550
weighted avg	0.96	0.96	0.96	550

FIG. 4: CLASSIFICATION REPORT OF LOGISTIC REGRESSION ALGORITHM

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	24
banana	1.00	1.00	1.00	19
blackgram	0.95	1.00	0.98	20
chickpea	1.00	1.00	1.00	26
coconut	1.00	1.00	1.00	27
coffee	1.00	1.00	1.00	26
cotton	0.96	1.00	0.98	24
grapes	1.00	1.00	1.00	23
jute	0.90	0.87	0.89	31
kidneybeans	0.96	1.00	0.98	25
lentil	0.96	1.00	0.98	26
maize	1.00	0.96	0.98	26
mango	1.00	1.00	1.00	31
mothbeans	1.00	0.92	0.96	25
mungbean	1.00	1.00	1.00	29
muskmelon	1.00	1.00	1.00	26
orange	1.00	1.00	1.00	34
papaya	1.00	1.00	1.00	23
pigeonpeas	1.00	0.96	0.98	24
pomegranate	1.00	1.00	1.00	19
rice	0.85	0.88	0.86	25
watermelon	1.00	1.00	1.00	17
accuracy			0.98	550
macro avg	0.98	0.98	0.98	550
weighted avg	0.98	0.98	0.98	550

FIG. 5: CLASSIFICATION REPORT OF K-NN ALGORITHM

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	24
banana	1.00	1.00	1.00	19
blackgram	1.00	1.00	1.00	20
chickpea	1.00	1.00	1.00	26
coconut	1.00	1.00	1.00	27
coffee	1.00	1.00	1.00	26
cotton	1.00	1.00	1.00	24
grapes	1.00	1.00	1.00	23
jute	0.91	0.97	0.94	31
kidneybeans	1.00	1.00	1.00	25
lentil	1.00	1.00	1.00	26
maize	1.00	1.00	1.00	26
mango	1.00	1.00	1.00	31
mothbeans	1.00	1.00	1.00	25
mungbean	1.00	1.00	1.00	29
muskmelon	1.00	1.00	1.00	26
orange	1.00	1.00	1.00	34
papaya	1.00	1.00	1.00	23
pigeonpeas	1.00	1.00	1.00	24
pomegranate	1.00	1.00	1.00	19
rice	0.96	0.88	0.92	25
watermelon	1.00	1.00	1.00	17
accuracy			0.99	550
macro avg	0.99	0.99	0.99	550
weighted avg	0.99	0.99	0.99	550

FIG. 6: CLASSIFICATION REPORT OF RANDOM FOREST ALGORITHM

TABLE 1: PERFORMANCE MEASURES OF MACHINE LEARNING ALGORITHMS

	Precision	**Recall**	**F1-score**	**Accuracy**
Decision Trees	0.80	0.86	0.81	85.27
Naïve Bayes	0.99	0.99	0.99	85.27
Logistic Regression	0.96	0.95	0.96	95.63
K-Nearest Neighbour	0.98	0.98	0.98	98.00
Random Forest	0.99	0.99	0.99	99.27

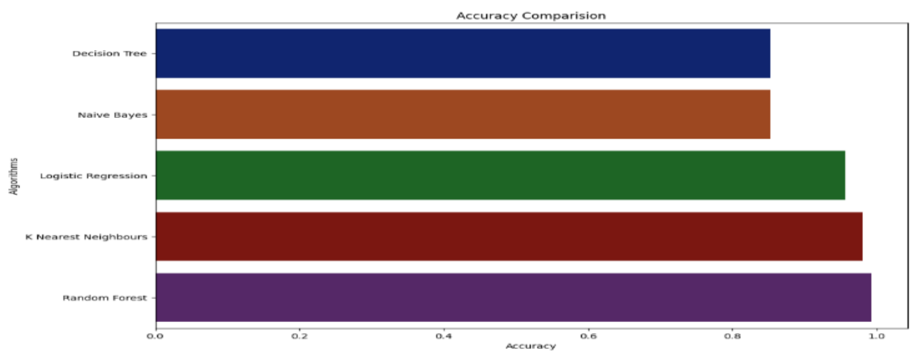

FIG. 7: ACCURACY COMPARISON CHART

V. Conclusion

The technique is extensive and uses soil factors such as N, P, K, Ph, humidity, rainfall, and temperature to suggest the best crops for a given set of soil conditions. It lowers the possibility of making poor decisions by helping farmers choose profitable crops, increasing total output. With an astounding 99.27% accuracy rate, the random forest method has the potential for much more improvement with IoT connection. Efficient agricultural operations are made possible by the incorporation of real-time soil data collected by deployed sensors, which improves accuracy and precision. As a result, this integrated system enables increased productivity, demonstrating the possibility for sustainable and advanced agriculture.

References

[1] Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., & Nisha, J. (2017, January). Crop recommendation system for precision agriculture. In *2016 eighth international conference on advanced computing (ICoAC)* (pp. 32-36). IEEE.

[2] Shalini, L., Manvi, S. S., Gardiner, B., & Gowda, N. C. (2022, December). Image Based Classification of COVID-19 Infection Using Ensemble of Machine Learning Classifiers and Deep Learning Techniques. In *2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)* (Vol. 1, pp. 1-6). IEEE.

[3] Kumar, R., Singh, M. P., Kumar, P., & Singh, J. P. (2015, May). Crop Selection Method to maximize crop yield rate using machine learning technique. In *2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM)* (pp. 138-145). IEEE.

[4] Rajak, R. K., Pawar, A., Pendke, M., Shinde, P., Rathod, S., & Devare, A. (2017). Crop recommendation system to maximize crop yield using machine learning technique. *International Research Journal of Engineering and Technology*, *4*(12), 950-953.

[5] Shalini, L., Manvi, S. S., Gowda, N. C., & Manasa, K. N. (2022, June). Detection of phishing emails using machine learning and deep learning. In *2022 7th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1237-1243). IEEE.

[6] Rekha, K. B., & Gowda, N. C. (2020, October). A framework for sentiment analysis in customer product reviews using machine learning. In *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)* (pp. 267-271). IEEE.

[7] Thomas, K. T., Varsha, S., Saji, M. M., Varghese, L., & Thomas, E. J. (2020). Crop prediction using machine learning. *International Journal of Future Generation Communication and Networking*, *13*(3), 1896-1901.

[8] Rajeswari, S. R., Khunteta, P., Kumar, S., Singh, A. R., & Pandey, V. (2019). Smart farming prediction using machine learning. *International Journal of Innovative Technology and Exploring Engineering*, *8*(07), 250-270.

[9] Jaiswal, S., Kharade, T., Kotambe, N., & Shinde, S. (2020). Collaborative recommendation system for agriculture sector. In *ITM web of conferences* (Vol. 32, p. 03034). EDP Sciences.

[10] Tonni, K. F., & Chowdhury, M. (2023). A machine learning based crop recommendation system and user-friendly android application for cultivation. *International Journal of Smart Technology and Learning*, *3*(2), 168-186.

[11] Anguraj, K., Thiyaneswaran, B., Megashree, G., Shri, J. P., Navya, S., & Jayanthi, J. (2021). Crop recommendation on analyzing soil using machine learning. *Turkish Journal of Computer and Mathematics Education*, *12*(6), 1784-1791.

[12] Yakoobi, M. M., & Gowda, N. C. (2023, May). Deep Learning-Based Solution for Differently-Abled Persons in The Society. In *2023 4th International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE.

[13] Waikar, V. C., Thorat, S. Y., Ghute, A. A., Rajput, P. P., & Shinde, M. S. (2020). Crop prediction based on soil classification using machine learning with classifier ensembling. *Int. Res. J. Eng. Technol*, *7*(5).

[14] Mariappan, A. K., Madhumitha, C., Nishitha, P., & Nivedhitha, S. (2020). Crop recommendation system through soil analysis using classification in machine learning. *International Journal of Advanced Science and Technology*, *29*(3), 12738-12747.

[15] Suruliandi, A., Mariammal, G., & Raja, S. P. (2021). Crop prediction based on soil and environmental characteristics using feature selection techniques. *Mathematical and Computer Modelling of Dynamical Systems*, *27*(1), 117-140.

[16] Nishant, P. S., Venkat, P. S., Avinash, B. L., & Jabber, B. (2020, June). Crop yield prediction based on Indian agriculture using machine learning. In *2020 international conference for emerging technology (INCET)* (pp. 1-4). IEEE.

[17] Gowda, N. C., & Malakreddy, B. (2023, February). A Trust Prediction Mechanism in Edge Communications using Optimized Support Vector Regression. In *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 784-789). IEEE.

[18] Doshi, Z., Nadkarni, S., Agrawal, R., & Shah, N. (2018, August). AgroConsultant: intelligent crop recommendation system using machine learning algorithms. In *2018 fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-6). IEEE.

[19] Mythili, K., & Rangaraj, R. (2021). Deep learning with particle swarm based hyper parameter tuning based crop recommendation for better crop yield for precision agriculture. *Indian Journal of Science and Technology*, *14*(17), 1325-1337.

[20] Akshatha, K. R., & Shreedhara, K. S. (2018). Implementation of machine learning algorithms for crop recommendation using precision agriculture. *International Journal of Research in Engineering, Science and Management (IJRESM)*, *1*(6), 58-60.

[21] Choudhury, S. S., Mohanty, S. N., & Jagadev, A. K. (2021). Multimodal trust based recommender system with machine learning approaches for movie recommendation. *International Journal of Information Technology*, *13*, 475-482.

[22] Majumdar, J., Naraseeyappa, S., & Ankalaki, S. (2017). Analysis of agriculture data using data mining techniques: application of big data. *Journal of Big data*, *4*(1), 20