



Intelligent Crop Recommender System for Yield Prediction Using Machine Learning Strategy

Atchukatla Maheswary¹ · Sanam Nagendram² · Kasi Uday Kiran³ ·
Shaik Hasane Ahammad³ · Putcha Poorna Priya⁴ · Md. Amzad Hossain⁵ ·
Ahmed Nabih Zaki Rashed^{6,7}

Received: 9 November 2023 / Accepted: 14 February 2024 / Published online: 14 March 2024
© The Institution of Engineers (India) 2024

Abstract For most developed nations, agriculture is a significant economic force. The realm of contemporary agriculture is consistently growing with evolving farming techniques and agricultural innovations. Farmers face challenges in keeping pace with the evolving demands of the planet and meeting the requirements of profitable initiatives, characters, and various other stakeholders. Climate change brought on by industry emissions and soil erosion, soil's nutrient deficiency due to mineral's absence, which results in reduced crop growth, and the cultivation of the same crops repeatedly without trying out new varieties are a few of the difficulties farmers face. Without considering the lower quality or quantity, they arbitrarily infuse fertilizers. Using two separate metrics, entropy and Gini indexes, the study analyzes well-known procedures with *K*-nearest neighbor (KNN), decision tree (DT), and random forest (RF) classifier practices. Moreover, the precision under the

agriculture paradigm, particularly “crop recommender systems,” includes these methods. Based on the outcomes, the random forest strategy outperforms the other approaches to model accuracy and reliability.

Keywords Crop yield prediction · Fertilizer · Random forest · Machine learning · Decision tree

Introduction

The importance of crop production forecasting is increasing as concerns over food security grow. Early crop yield prediction plays a significant role in averting famine by anticipating the amount of food available for the expanding global population. Increased crop yield output is a workable answer to one of the world's most serious

✉ Shaik Hasane Ahammad
ahammadklu@gmail.com

✉ Md. Amzad Hossain
mahossain.eee@gmail.com

✉ Ahmed Nabih Zaki Rashed
ahmed_733@yahoo.com

Atchukatla Maheswary
amaheswary@svcet.in

Sanam Nagendram
reena1286@gmail.com

Kasi Uday Kiran
uk_ee@kluniversity.in

Putcha Poorna Priya
putchapoornapriya@gmail.com

¹ Department of Electronics and Communication Engineering,
Sri Venkateswara College of Engineering & Technology,
Thirupatchur, India

² Department of Artificial Intelligence, KKR & KSR Institute
of Technology and Sciences, Guntur, India

³ Department of Electronics and Communication
Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur 522502, India

⁴ Department of Electronics and Communication Engineering,
Dadi Institute of Engineering and Technology College,
Anakapalle, India

⁵ Department of Electrical and Electronic Engineering,
Jashore University of Science and Technology, Jashore 7408,
Bangladesh

⁶ Electronics and Electrical Communications Engineering
Department, Faculty of Electronic Engineering, Menoufia
University, Menouf 32951, Egypt

⁷ Department of VLSI Microelectronics, Saveetha School
of Engineering, Saveetha Institute of Medical and Technical
Sciences (SIMATS), Saveetha University, Chennai,
Tamilnadu, India

problems: hunger [1]. There is still not enough food available for 820 million people worldwide, based on the World Health Organization (WHO) analysis. By 2030, as stated by the United Nations' Sustainable Development Goals (UNSDGs), objective is to end starvation, accomplish food protection, and support reliable and sustainable agricultural science. By 2050, the world's population of 9.3 billion will require 60% more food than the current production, according to the Food and Agricultural Organization (FAO) [2, 3]. Crop production prediction can provide essential information to attain the goal and end hunger. However, many considerations influence productivity in the crop division, generating the future challenge of developing a consistent model for forecasting employing traditional procedures. Moreover, the prediction of yield for the novel design and training process should be associated with the developments of advanced computer technology. A further factor is that 51% of farmers wish to relocate to cities rather than pass on their agriculture to the next generation [4]. Finally, it is because farmers frequently choose the wrong crops, associated with crop selection that would not produce the soil or planting during the stormy season. The decision may have been made without prior knowledge because the farmer may have bought the land from someone else. Less yield will always result from choosing the incorrect crop. It is exceedingly challenging to exist if the family ultimately depends on this revenue. Machine learning is a significant resource for deciding crop growing season activities, crop types to be sown, and agricultural yield predictions [5]. Many practices are associated with machine learning strategies for the prediction enhancement of crop recommendation systems. Indeed, methods linked with supervised machine learning techniques are employed for industrial purposes, from analyzing store patron behavior to forecasting phone usage. Besides agricultural applications also, these strategies are applied [6]. Alongside, crop forecasting remains agriculture's crucial

and most formidable problem, and numerous standards are designed combined with a tested case up to this point. This challenge demands the use of many datasets due to the complexity of the factors that affect crop yield, including atmospheric conditions, fertilizer type, soil, and seed. Farmers receive aid for opting the crop through the application of machine learning (ML) and data mining tools, taking into account factors like soil qualities, distinctive terrestrial regions, sowing season, and ecological considerations. Hence, it can be inferred that estimating agricultural productivity requires many intricate steps and is not a straightforward operation [7].

Figure 1 illustrates the recommendation system block representation. These algorithms' statistical characteristics can result in a sizable boost in yield. It is desired to have a high degree of precision, considering the inability to accomplish this will have significant damaging consequences, including seed and productivity losses and a substantial decline in profitability. Although prediction strategies of crop yield may currently reasonably estimate within the original yield base fetching over the performance outcome with the system, which is preferred [8]. Prospective investigators face obstacles in conducting country case findings due to challenges in accessing accurate and up-to-date information.

Main contribution of the research is to suggest the intelligent crop recommendation system for better yield prediction. Early crop yield prediction plays a significant role in averting famine by anticipating the amount of food available for the expanding global population. To predict the crop recommendation that would produce effective procedure in the actual land site, it employs three distinctive supervised-based machine learning prototypes: decision tree, random forest, and KNN. Consequently, the comparison is also conducted for the proposed methods to know the performance analysis and to conclude, with two separate metrics, entropy and Gini indexes.

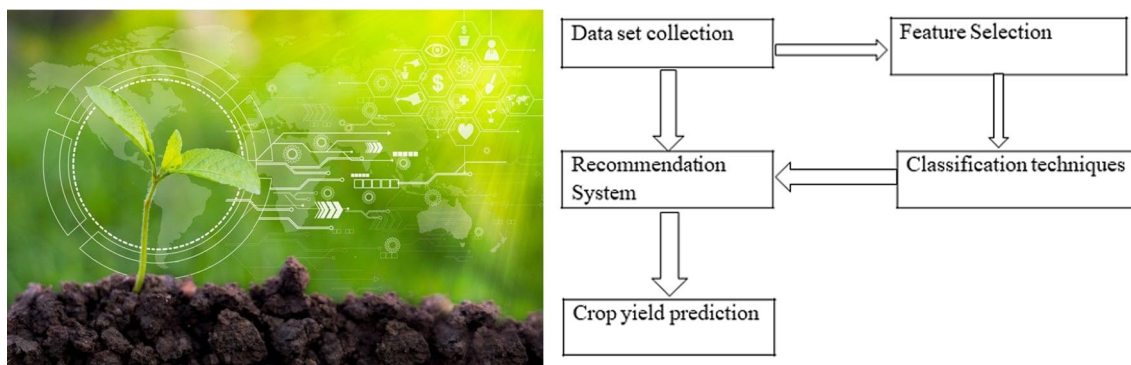


Fig. 1 Block representation of the recommendation system

Literature Survey

Several ideas have been put out in the past to increase crop prediction accuracy considering the importance of this field. Based on soil factors (potassium, Ph, and nitrogen) combined with constraints associated with the environment, as employed the feed-forward back propagation (FFBP)-based artificial neural network (ANN) technology for forecasting and modeling the crop yields in rural regions (rainfall, humidity, etc.) [9]. Many variables, including rainfall, groundwater, the area under cultivation, and the kind of soil, were considered in the analysis to ensure the highest crop productivity. The study examined three distinct strategical types for classification purposes that include fuzzy logic, SVM, and KNN, as well as modified KNN (MKNN). Whereas, for the clustering process, the *K*-means technique is employed [10]. MKNN provided with the good quality prediction result among the three algorithms following the analysis. Besides, the feasible development of the tool for the cultivators associated with the application of various challenges as the current industrialized world is facing [11]. Moreover, the farmers or cultivators utilize such advanced computer technology for the individual execution of the process and undergoing various investigational tests through the information entered such as location and time (Years) for the opted crop [12]. The user can select a procedure that determines and evaluates the actual outcome for the best input application within the system to recommend the user. The rate of crop’s yield is visible in the productions, which are derived from datasets containing past database outcomes adapted into a consistent organization. Naive Bayes and KNN are the utilized machine learning models [13].

Data on crops spanning the last decade were collected for the dataset, sourced from various locations, containing authority sites. For humidity and temperature analysis, the electronic component utilized is soil and DHT11 sensors. Whereas, for the purpose of data gathering with the atmospheric analysis, the Arduino-Uno board integrated to Atmega CPU with collaboration of things (IoT) can be used.

Lekhaa et al. [14], Lobell and Burke [15], boosting technique, that employments frail rules through an reiterative method to transport augmented precision, remains for enhancement of Naive Bayes (NB) strategy, which is a supervised-based learning procedures attained model precision of about 96% as mentioned [16]. Operating with the stacking development under regression analysis for the improved estimate and forecast enhances the three regression techniques. The suggested approach within the system currently associated through the random forest, and Naive Bayes correspondingly were compared for the performance analysis [17]. Nevertheless, the model accuracy is bagging when compared to the random forest based on the probability function built with the system. Crop production is

forecasting using several machine-based approaches with accuracy and rate of error comparison under specific areas [18]. Employing available reserves, an organization is introduced to address this challenge using predictions on crop sustainability. This system also offers recommendations based on ML models established, considering crucial conservation and commercial influences [19, 20].

Table 1 lists the existing crop prediction approaches. Using 22 various sorts of featuring dataset within the crop variants, the study compares different supervised-based-learning practices that include decision tree, random forest, and KNN. Nonetheless, for the analysis, there are criteria such as the Gini index and entropy employed techniques of random forest and decision tree classifiers [10]. Based on precision, recall, and *F1* score, the results demonstrate that the suggested machine learning technique performs as well as the best accuracy [5, 26]. Several machine learning algorithms, including KNN, support vector machine (SVM), Fuzzy logic (FL), artificial neural network (ANN), multiple linear regression (MLR), and random forest (RF) strategies, have been deployed on the provided datasets and evaluated for yield to accuracy. Comparisons are made between the algorithms’ accuracy [11] by Sunil et al. [19] and Van Klompenburg et al. [27]. The results demonstrate that random forest decision tree classifier, with a 99.40% accuracy rate, remains exact precision within the conventional strategies as deployed in the datasets provided.

Datasets and Methodology

Data accumulated and filtered from various resources are subjected to data collection. First, tracking past occurrences is kept safe for accurate data analysis for repetitive pattern detection. Then, the raw data are modified to a form for the functioning of MI techniques to provide better insights under the data pre-processing. Consequently, the splitting, fitting, and process of checking the range over the dataset trained, within the model prediction, thus generating the classification and confusion matrix report undergo.

Table 2 demonstrates the sample dataset fed to the methodology for executing the recommendation process.

Table 1 Existing crop prediction approaches

References	Model	Accuracy (%)
Portugal et al. [21]	Naïve Bayes	97
Priya et al. [22]	SVM, Naïve Bayes, and random forest classifier	95
Pudumalar et al. [23]	<i>K</i> -nearest neighbor	93
Rajak et al. [24]	Linear regression	92
Rale et al. [25]	Linear regression	91.05

Table 2 Sample dataset

	P	N	K	Humidity	Temperature	Rainfall	pH	Label
0	38	85	50	79.356595	19.57849758	121.89525	7.698225	rice
1	45	78	39	80.232659	21.59898974	158.59256	5.565254	rice
2	48	90	43	75.548445	22.35056591	160.59595	6.584224	rice
3	38	76	52	72.648265	20.55494950	203.59626	6.123575	rice
4	440	69	50	65.205493	21.28945626	220.79812	6.321024	rice
...
3024	39	99	35	80.125389	20.65949625	200.59795	6.795252	Corn
3025	22	105	29	54.489595	26.55959252	210.59425	7.123589	Corn

From the Kaggle website, the data are collected for the “crop recommendation”. Moreover, the considered terms within the dataset of the entire 22 variants of crops linked to features and labels are phosphorus content ratio (P), temperature (°C), relative humidity %, nitrogen content ratio (N), rainfall (mtrs), pH value, and potassium content ratio (K).

Figure 2 depicts the steps for methodology. Besides, the sampling frequency ranging within the information set up to the actual and prediction will be assigned by the confusion matrix as false positive, false negative, and true positive, true negative. Furthermore, the statistical analysis developed within the assessment for the classification process developed for the prediction practices; thus, classification report is significant. It provides three results of the performance metrics. The capacity of a classifier to count the prediction count estimate linked for the significant analysis in the outcome that usually signifies the accuracy of the model termed as the rate of precision. The proportion for a true and false positive is computed for each class. Nevertheless, the classifier’s recall determines the capacity for positive extraction built within the development of the confusion matrix. The proportion for true positives and false negatives is computed for individually. Moreover, the *F1* score, with 0.0 being the poorest and 1.0 being the greatest, is a weighted harmonic-mean of recall and precision. *F1* scores are frequently lower than accuracy assessments because recall and precision rates are factored in during computations. Furthermore, the prediction rates

under the ratio of prediction count signified model accuracy. Figure 3 shows the confusion matrix.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$F1 \text{ score} = \frac{2 \times \text{Precision}}{(\text{Precision} + \text{Recall})} \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

here TP–true positive, TN–true negative, FP–false positive, and FN–false negative.

Proposed Study

K-nearest Neighbor Classifier

A straightforward machine learning process, the KNN relies on the supervised-based learning methodology [2]. There are two features of the KNN method. Figure 4 describes the practical steps for KNN. Initially, the distribution parameters do not need to be determined because of the dependency of the dataset on the model and, as a result, subjected as non-parametric. Secondly, it simply stores the training data; no

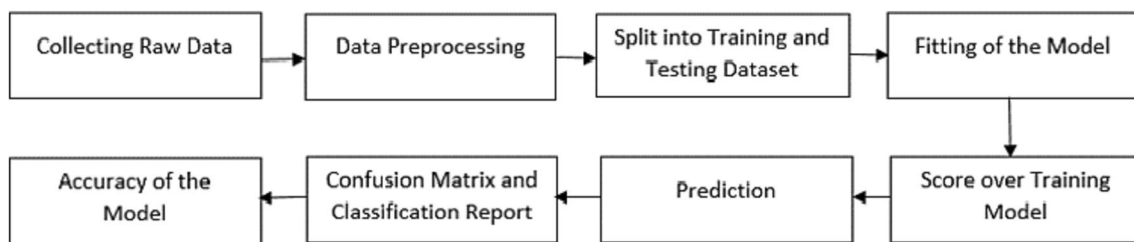


Fig. 2 Methodology steps

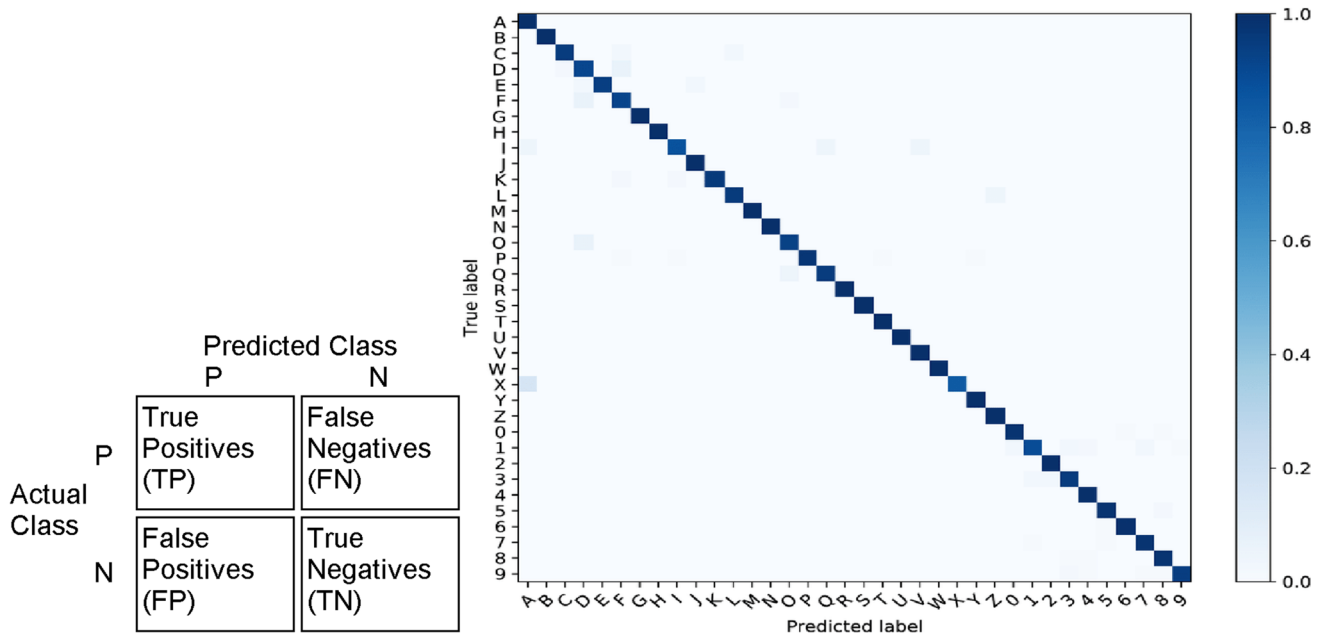


Fig. 3 Confusion matrix

learning is happening. Finally, the testing phase includes the classification of the dataset, which requires a lot of memory and increases testing time.

$$\text{Euclidean distance} = d(b, a) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \tag{5}$$

KNN outperforms the supervised learning within the test case of the method fed to the system. For the quality process, it is termed as a lazy learner.

Decision Tree Classifier

Another supervised-based learning is the decision tree, which can be deployed for the case of classification analysis and regression approaches within the individual path associated with subject-leading decisions within the class [6, 12]. Nevertheless, the estimate decision is developed through the instance comparison over with the decision can be attributed to the outcome of split and jump for upcoming node region. Splitting continues, producing the sub-trees until the leaf node reaches. Thus, class labels were determined instantaneously. Therefore, recursive portioning occurs. The stepwise representation of the pictorial format for the decision tree strategy is shown in Fig. 5.

Entropy and Gini indexes and information gain are determined for every split. Furthermore, the information gain

signifies the measure in the rate of information acquired in the generated attribute factor and thus entropy minimization. Moreover, the degree of impurity within the given attribute is indicated by entropy and by subtracting the overall quantities of the squared term of probabilities within the domain of class counted with 1, consequently determining the Gini index of coefficient.

$$I(\text{Attribute}) = \frac{\sum y_i + n_i}{y + n} \tag{6}$$

Here $I(\text{Attribute})$ —Nominal information, n_i —no values, and y_i —yes values for attributes, y & n are yes, no for overall sample counts.

$$\text{Information Gain} = \text{Entropy}(S) - I(\text{Attribute}) \tag{7}$$

$$\text{Entropy}(S) = -H(\text{yes}) * \log_2 H(\text{yes}) - H(\text{no}) * \log_2 H(\text{no}) \tag{8}$$

$$\text{Entropy}(S) = \sum_{i=1}^N -h_i \log_2 h_i \tag{9}$$

$$\text{Gini Index} = 1 - \sum_{i=1}^c (H_i^2) \tag{10}$$

here S —sample, $H(\text{yes})$ and $H(\text{no})$ —probability chance of yes.

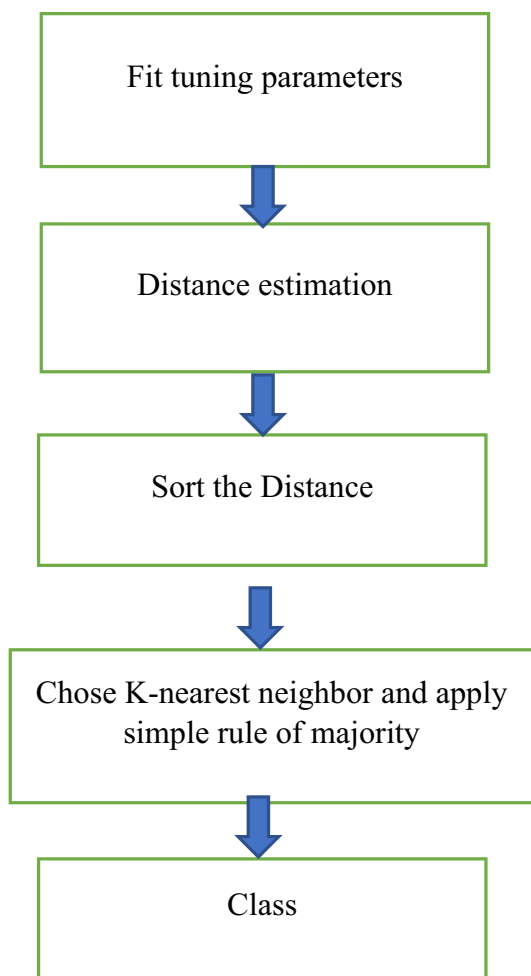


Fig. 4 KNN classifier steps

Random Forest Classifier

This approach entails multiple DT classifiers for the simulation’s execution enhancement. It uses ensemble-based-knowledge theory [7, 21]. Randomly, decision trees are created within the training data to predict outcomes. The decision of the final prediction is made through a majority voting process.

Figure 6 demonstrates the strategy for the random forest classifier to apply with the recommendation system. At the current step, each estimator gives the determined prediction; thus, the category with the highest votes is allocated.

Results and Discussions

In the model of KNN, the accuracy is dependent on the neighbors. Here, $K=5$. From the figure, the significance of linear or straight-line is for the demonstration of prediction accuracy and that with the outer projections determining the case of lousy prediction estimates.

Figure 7 shows the analysis for the prediction of crop recommendation through the KNN approach. Within the selected criterion of Gini and entropy for the samples attributed to the average count of the information fed to the system at the maximum probability distribution.

Figure 8 demonstrates the result analysis for the two criteria chosen for the classifier approach in the recommendation system. The decision tree with a measure of Gini showed a minimal count of dice projections associated to the entropy criterion. Thus, Gini outperforms the entropy under accuracy.

Fig. 5 Decision tree strategy

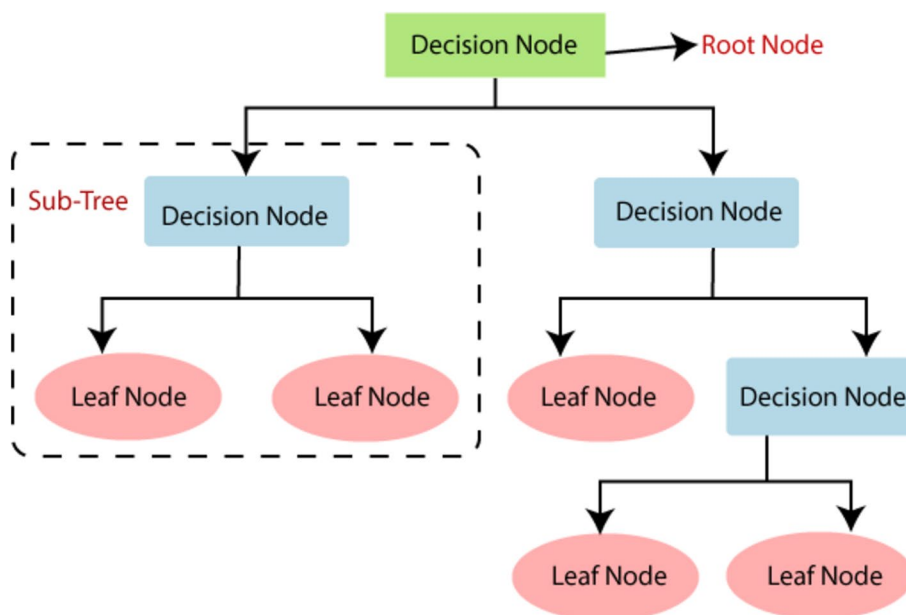


Fig. 6 Random forests classifier strategy

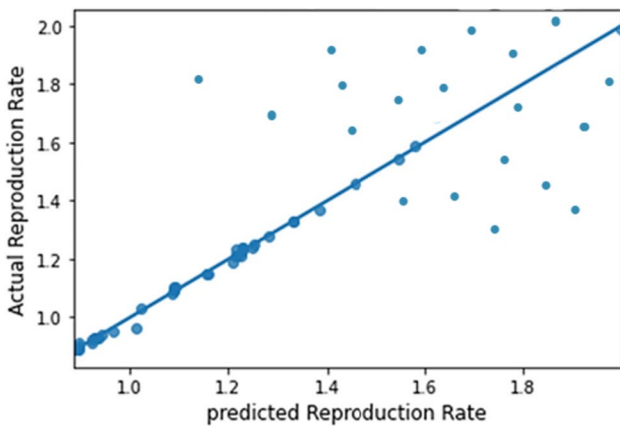
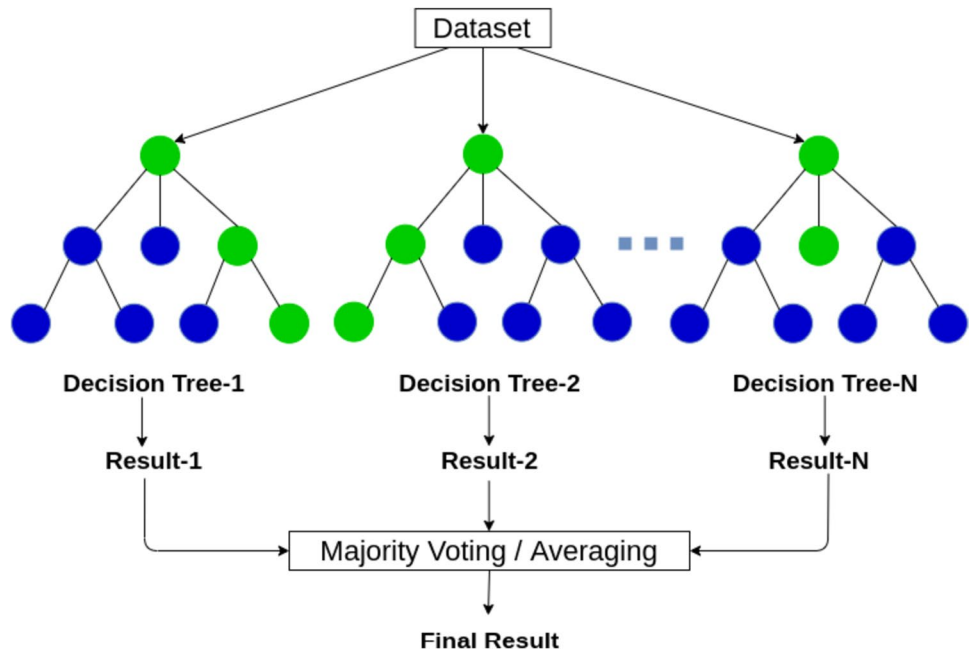


Fig. 7 KNN prediction

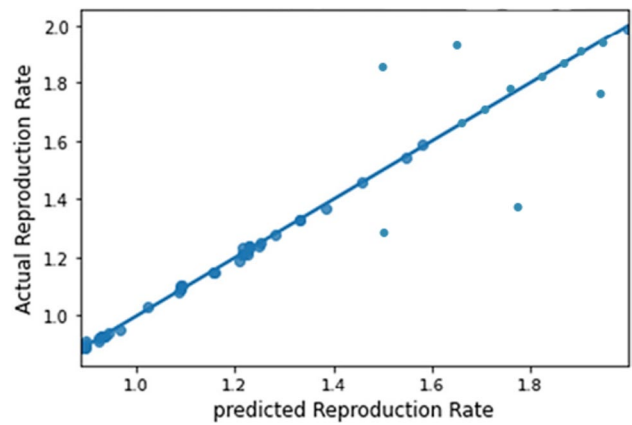


Fig. 9 Decision tree prediction

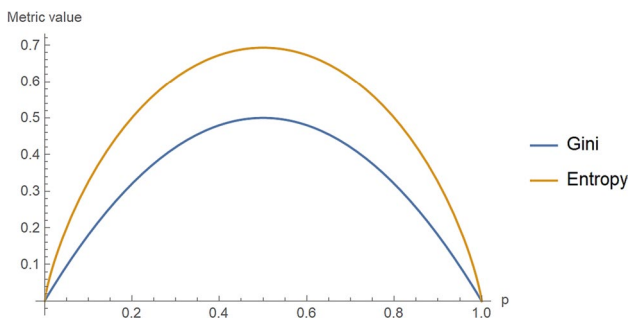


Fig. 8 Analysis of Gini and entropy criterion's predictions

Figure 9 illustrates the decision tree strategy's prediction outcome compared with the actual and prediction case under the criteria linked to the system. While designing with the random forest strategy, the estimator count is set to 100, i.e., 100 decision trees. Alongside, the estimator count does not affect the model accuracy and thus accomplishes the best outcome. Random forest demonstrated equivalent predictions under both criteria.

Figure 10 shows the prediction outcome for the strategy of the random forest classifier compared with the actual and prediction case under the criteria linked to the sustain. The random forest classifier.

Figure 11 illustrates the model accuracy comparison for the proposed methods in terms of their model accuracy, which is analyzed and tabulated with its training score fed

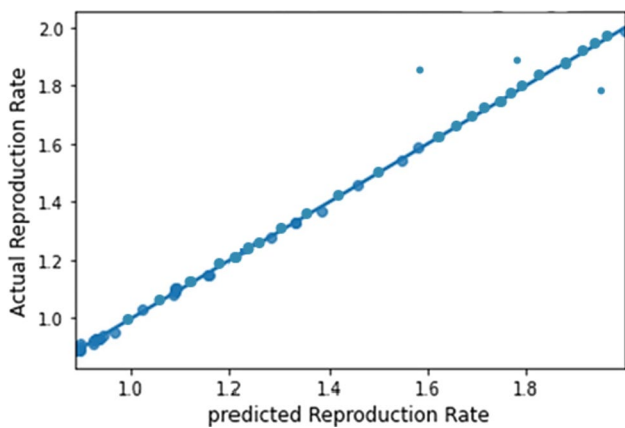


Fig. 10 Random forest prediction

with the recommendation system. Table 3 signifies the model accuracy and training scores comparison for the proposed methods in processing the recommendation system for the crop.

Conclusion

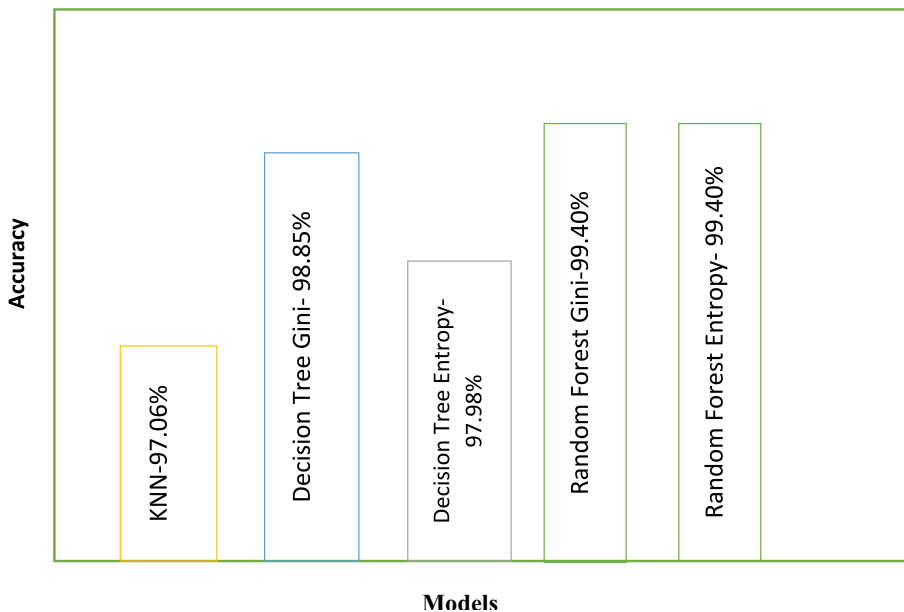
The study introduces an efficient crop recommendation system using classifier models. To predict the crop recommendation that would produce effective procedure in the actual land site, it employs three distinctive

supervised-based ML prototypes: decision tree, random forest, and KNN. Consequently, the comparison is also conducted for the proposed methods to know the performance analysis. To conclude, with two separate metrics, entropy and Gini Indexes, the study analyzes well-known procedures with *K*-nearest neighbor (KNN), random forest classifier, and decision tree practices. It is evident from the outcomes of crop prediction; the random forest classifier outperformed the other strategies in Gini- and entropy-based criteria with a greater accuracy rate of 99.04%. Contrarily, within the comparison, *K*-nearest neighbor attained the lowest accuracy (97.06%) and decision tree classifier (~98%). Regarding accuracy, the decision tree Gini criterion outperformed the decision tree entropy criterion with a score of 98.85%.

Table 3 Model accuracy and training scores comparison for the proposed methods

Models	Accuracy	Training score
KNN	97.06	98.79
Decision tree with entropy	97.98%	100
Decision tree with Gini	98.85%	100
Random forest with entropy	99.04%	100
Random forest with Gini	99.04%	100
Linear regression model	88.25%	–
Naive Bayes	82%	–

Fig. 11 Model accuracy of proposed methods



Authors Contributions Authors contributed equally in this work.

Funding Not applicable.

Data Availability Not applicable.

Declarations

Conflict of interest No competing interests.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

References

1. N. Balakrishnan, G. Muthukumarasamy, Crop production-ensemble machine learning model for prediction. *Int. J. Comput. Sci. Softw. Eng.* **5**(7), 148 (2016)
2. S. Bang, R. Bishnoi, A.S. Chauhan, A.K. Dixit, I. Chawla, Fuzzy logic based crop yield prediction using temperature and rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1–6). (IEEE, 2019).
3. J. Bobadilla, F. Ortega, A. Hernando, J. Alcalá, Improving collaborative filtering recommender system results and performance using genetic algorithms. *Knowl.-Based Syst.-Based Syst.* **24**(8), 1310–1316 (2011)
4. R. Burke, Hybrid recommender systems: survey and experiments. *User Model. User-Adap. Interact.* **12**, 331–370 (2002)
5. A. Chlingaryan, S. Sukkarieh, B. Whelan, Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. *Comput. Electron. Agric.* **151**, 61–69 (2018)
6. V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, A.P. Janani, An effective crop prediction using random forest algorithm. In 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1–5). (IEEE, 2020).
7. P. Gulati, S.K. Jha, Efficient crop yield prediction in India using machine learning techniques. *Int. J. Eng. Res. Technol. (IJERT) ENCADEMS–2020*, **8**(10) (2020).
8. A. Gupta, D. Nagda, P. Nikhare, A. Sandbhor, Smart crop prediction using IoT and machine learning. *Int. J. Eng. Res. Technol. (IJERT)*, 2278–0181 (2021).
9. L.K. Johnson, J.D. Bloom, R.D. Dunning, C.C. Gunter, M.D. Boyette, N.G. Creamer, Farmer harvest decisions and vegetable loss in primary production. *Agric. Syst.* **176**, 102672 (2019)
10. J.W. Jones, J.M. Antle, B. Basso, K.J. Boote, R.T. Conant, I. Foster, T.R. Wheeler, Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. *Agric. Syst.* **155**, 269–288 (2017)
11. M. Kalimuthu, P. Vaishnavi, M. Kishore, Crop prediction using machine learning. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 926–932). (IEEE, 2020).
12. S. Kulkarni, S.N. Mandal, G.S. Sharma, M.R. Mundada, Predictive analysis to improve crop yield using a neural network model. In 2018 International Conference on Advances in Computing, Communications, and Informatics (ICACCI) (pp. 74–79). (IEEE, 2018).
13. H. Lee, A. Moon, Development of yield prediction system based on real-time agricultural meteorological information. In 16th International Conference on Advanced Communication Technology (pp. 1292–1295). (IEEE, 2014)
14. T.R. Lekhaa, Efficient crop yield and pesticide prediction for improving agricultural economy using data mining techniques. *Int. J. Mod. Trends Eng. Sci. (IJMTES)* **3**(10), 11–28 (2016)
15. D.B. Lobell, M.B. Burke, On the use of statistical models to predict crop yield responses to climate change. *Agric. For. Meteorol. Meteorol.* **150**(11), 1443–1452 (2010)
16. P. Malik, S. Sengupta, J.S. Jadon, Comparative analysis of soil properties to predict fertility and crop yield using machine learning algorithms. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 1004–1007). (IEEE, 2021)
17. K. Nischitha, D. Vishwakarma, M.N. Ashwini, M.R. Manjuraju, Crop prediction using machine learning approaches. *Int. J. Eng. Res. Technol. (IJERT)* **9**(08), 23–26 (2020)
18. S.M. Pande, P.K. Ramesh, A. Anmol, B.R. Aishwarya, K. Rohilla, K. Shaurya, Crop recommender system using machine learning approach. In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1066–1071). (IEEE, 2021).
19. G.L. Sunil, V. Nagaveni, U. Shruthi, A review on prediction of crop yield using machine learning techniques. In 2022 IEEE Region 10 Symposium (TENSYMP) (pp. 1–5). (IEEE, 2022)
20. A. Poriya, T. Bhagat, N. Patel, R. Sharma, Non-personalized recommender systems and user-based collaborative recommender systems. *Int. J. Appl. Inf. Syst.* **6**(9), 22–27 (2014)
21. I. Portugal, P. Alencar, D. Cowan, The use of machine learning algorithms in recommender systems: a systematic review. *Expert Syst. Appl.* **97**, 205–227 (2018)
22. R. Priya, D. Ramesh, E. Khosla, Crop prediction on the region belts of India: a Naïve Bayes MapReduce precision agricultural model. In 2018 International Conference on Advances in Computing, Communications, and Informatics (ICACCI) (pp. 99–104). (IEEE, 2018).
23. S. Pudumalar, E. Ramanujam, R.H. Rajashree, C. Kavya, T. Kiruthika, J. Nisha, Crop recommendation system for precision agriculture. In 2016 Eighth International Conference on Advanced Computing (ICoAC) (pp. 32–36). (IEEE, 2017)
24. R.K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, A. Devare, Crop recommendation system to maximize crop yield using machine learning technique. *Int. Res. J. Eng. Technol.* **4**(12), 950–953 (2017)
25. N. Rale, R. Solanki, D. Bein, J. Andro-Vasko, W. Bein, Prediction of crop cultivation. In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0227–0232). (IEEE, 2019)
26. M.T. Shakoor, K. Rahman, S.N. Rayta, A. Chakrabarty, Agricultural production output prediction using supervised machine learning techniques. In 2017 1st international conference on next generation computing applications (NextComp) (pp. 182–187). (IEEE, 2017).
27. T. Van Klompenburg, A. Kassahun, C. Catal, Crop yield prediction using machine learning: a systematic literature review. *Comput. Electron. Agric. Electron. Agric.* **177**, 105709 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.