Heliyon 10 (2024) e25112

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables

Biplob Dey [a,b], Jannatul Ferdous [a], Romel Ahmed [a,b,*]

[a] Department of Forestry and Environmental Science, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh
[b] Center for Research in Environment, iGen and Livelihood (CREGL), Sylhet 3114, Bangladesh

ARTICLE INFO

Keywords:
Crop recommendation
NPK
Machine learning
Agricultural crops
AI
Horticultural crops

ABSTRACT

Machine learning (ML) can make use of agricultural data related to crop yield under varying soil nutrient levels, and climatic fluctuations to suggest appropriate crops or supplementary nutrients to achieve the highest possible production. The aim of this study was to evaluate the efficacy of five distinct ML models for a dataset sourced from the Kaggle repository to generate practical recommendations for crop selection or determination of required nutrient(s) in a given site. The datasets contain information on NPK, soil pH, and three climatic variables: temperature, rainfall, and humidity. The models namely Support vector machine, XGBoost, Random forest, KNN, and Decision Tree were trained using yields of individual data sets of 11 agricultural and 10 horticultural crops, as well as combined yield of both agri-horticultural crops. The results strongly suggest to evaluate individual data sets separately for each crop category rather than using combined the data sets of both categories for better predictions. Comparing the five ML models, the XGBoost demonstrated the highest level of accuracy. The precision rates of XGBoost for recommending agricultural crops, horticultural crops, and a combination of both were 99.09 % (AUC 1.0), 99.3 % (AUC 1.0), and 98.51 % (AUC 0.99), respectively. This non-intrusive method for generating crop recommendations in diverse environmental conditions holds the potential to provide valuable insights for the development of a user-friendly AI cloud-based interface. Such an interface would enable rapid decision-making for optimal fertilizer applications and the selection of suitable crops for cultivation at specific sites.

1. Introduction

The economy of South Asia heavily relies on agriculture [1], which plays a vital role in ensuring both human survival and economic growth [2,3]. For over 13,000 years, people have been relying on it as their primary source of sustenance, and it remains the main food source for the world's population [4]. According to a projection, the current global population of 7.8 billion is projected to increase to 9.8 billion by 2050 [5], among them over 800 million individuals still lack sufficient access to food, and approximately 10 % of food production is lost due to factors such as pests, diseases, and adverse weather conditions [6]. These circumstances resulted in food insecurity, one of the biggest challenges that humanity is currently confronting in the 21st century [7]. To meet the growing global

* Corresponding author. Department of Forestry and Environmental Science, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh.

E-mail address: romel-fes@sust.edu (R. Ahmed).

https://doi.org/10.1016/j.heliyon.2024.e25112
Received 29 June 2023; Received in revised form 7 December 2023; Accepted 20 January 2024
Available online 26 January 2024

demand for food and ensure food security, it is crucial to adopt advanced agricultural practices and innovative technologies in farming.

The agricultural production heavily relies on local weather and climate conditions, as well as the occurrence and duration of extreme weather events [8,9], when making decisions about crop management. Inadequate availability of soil nutrients can have a negative impact on the biochemical and physiological performance of plants, ultimately resulting in reduced crop yields [10]. Maintaining soil productivity over the long term requires a balanced use of both organic and inorganic fertilizers, particularly in cases where the organic matter content is relatively low [11]. The fertility of the soil, growth of crops, and sustainable yield are all influenced by the kind and amount of fertilizers used, whether they are organic, inorganic, or a mixture of both. It is essential to regulate soil fertility in a logical manner and supplement any lacking nutrients through external fertilization to achieve a fruitful crop harvest, as suggested by sources [10,12].

Farmers often fail to consider the suitability of a site for a particular crop, leading to the planting of crops in unsuitable areas that resulted in the loss of productivity [9]. Hence, there is a need for an effective and highly precise system to suggest which crops to plant. The utilization of advanced technologies or automated farming techniques can be beneficial in addressing concerns related to environmental sustainability, soil optimization, and the collection and analysis of multiple heterogeneous variables. The potential to enhance crop cultivation methods based on available resources, adaptability to specific environments, and increased productivity is a significant area of interest. Having an accurate and efficient approach to crop management is essential in guaranteeing sustainable production and food security.

In recent years, there has been significant growth in predicting crop yields [13–17], identifying crop types [18,19], crop diseases identification [10,20–24], crop health assessment [25–27] through the application of ML and deep convolutional neural networks. An intelligent system called AgroConsultant was created by Doshi [28] to help farmers make informed decisions about which crop to cultivate based on their current environmental conditions such as soil type, thickness, pH, rainfall, and temperature. Waikar [29] used various ML algorithms such as Support vector machine (SVM), AdaBoost, Bagged tree, Artificial neural network (ANN), and Naïve Bayes to forecast crop yield based on soil types, soil properties and rainfall. Supervised ML algorithms, such as RF and SVM, were employed by Dubois [30] to develop models for predicting soil water potential in potato farming. Likewise ML algorithms were used by Ahmed and Hossain [31] to forecast wheat yield and multilayer perceptron (MLP), decision table, JRip algorithms (IoT framework) were employed by Gutiérrez [32] to suggest crop strategies techniques. Rajak [33] also used a soil database collected from farms and a dataset from a soil testing laboratory to recommend a crop based on site-specific parameters using SVM and ANN as machine learning models. The recent suggest that the advanced development of ML approaches can easily handle big data and precisely predict several cropping strategies, reducing the driving force of food security. Recent studies suggest that the advanced development of machine learning techniques is necessary to effectively manage large amounts of data and precisely predict various agricultural methods.

Despite the progress in the ML in predicting cropping strategies in smart farming, there are some pitfalls in model training of the data where necessary parameters were not taken into consideration. In previous studies [28,34–36], recommendations for crop selection have often utilized a combination of horticultural and agricultural crops for training models, rather than focusing solely on individual crop categories. Some research has concentrated on predicting crop yield by examining specific soil properties such as soil type, depth, pH, and topsoil thickness [28,37], while others have solely analyzed soil NPK levels and p^H [38]. In actuality, several interconnected and interdependent elements work in conjunction to dictate the crop's output. In order to accurately predict crop growth under specific weather conditions, it is essential to gather a wide range of data from different fields and geographic regions. This is because the nutrient and abiotic needs of crops vary depending on the type of crop and its location. Additionally, it remains uncertain how the combination of soil NPK content, soil pH, and climatic factors would impact the precision, accuracy, recall, and F1 score when using a single crop category of either agricultural or horticultural crops or combination of both. To address this, the present study employed a vast public dataset that considers the soil pH, external application of NPK fertilizer, and environmental factors such as rainfall, temperature, and humidity on a broad spectrum of crops. The data collected from a wide geographical region of India were sourced from Kaggle repository to evaluate the efficacy of ML algorithms in making efficient crop recommendations and to suggest the complementary correction measures to maximize the yield. This approach can be applied for various types of crops and regions having the similar environmental conditions.

2. Methodology

2.1. Data

The dataset utilized in this research was sourced from the Kaggle repository [39], which was collected over a period of time by the Indian Chamber of Food and Agriculture [40]. It comprises a total of 2100 data points, encompassing 11 agricultural crops and 10 horticulture crops grown under variables for NPK fertilizer, soil pH, and climatic factors such as rainfall, temperature, and humidity. The scientific names of both agricultural and horticultural crops are reported in the supplementary file (S1). The dataset shows the mean values of externally applied N, P, and K fertilizers for agricultural crops were 56, 52.11, and 31.64 kg/hector, respectively, and corresponding environmental conditions such as 24.89 °C± 4.02 °C temperature, 64.20 ± 24.10 % relative humidity (RH), 6.67 ± 0.85 pH, and 90.90 ± 61.64 mm rainfall. On the other hand, horticultural crops were found to have received an average of 47.52 kg/hector of N fertilizer, 53.29 kg/hector of P fertilizer, and 69.09 kg/hector of K fertilizer. For horticultural crops, the temperature recorded was 26.19 °C± 5.83 °C, RH was 81.82 ± 14.8 %, pH was 6.31 ± 0.57, and rainfall was 112.67 ± 43.47 mm. In order to evaluate the efficacy of the models, the crops were divided into three distinct categories: agricultural crops (AC), horticultural crops (HC), and a combination of agri-horticultural crops terming them hereafter as AC-model, HC-model and mixed crop model (Co) respectively. This dataset possesses a unique quality as it encompasses diverse geographical conditions and a wide range of crops, thereby exhibiting the potential to be utilized in various regions across the world having similar environmental conditions.

2.2. Prediction of crop using ML techniques

The study aimed to predict the selection of crops based on several factors, including NPK fertilizer, soil pH, and climatic factors, using regression algorithms (Fig. 1). To accomplish this objective, five machine learning algorithms were utilized including SVM, random forest (RF), eXtreme gradient boosting (XGBoost), K-nearest neighbors (KNN), and decision tree (DT).

2.2.1. Decision tree

A decision tree is a method for categorizing crops based on attribute-value tests related to factors like NPK fertilizer, soil pH, and climate. These tests are used to create the tree, starting with a training set of data points, each with attribute values and a corresponding crop category. The choice of which attribute to test is based on its ability to differentiate between crop categories, and this process is recursive, leading to various tree sizes. Pruning can be applied to prevent overfitting and improve the tree's classification performance in terms of portability and generalization. GridSearchCV was used to fine-tuning and finding the optimal hyperparameters for the decision tree model. Specifically, max features = 'auto', the best estimator, ccp alpha = 0.001, criterion = 'entropy', random state = 2, and max depth = 5 were employed in the training of the decision tree model.

2.2.2. eXtreme gradient boosting

XGBoost is widely regarded as an advancement in machine learning algorithms, owing to its integration of a gradient-boosted decision tree. This feature endows the model with superior flexibility, speed, and performance compared to other models. Decision tree ensembles are built using numerous decision tree models. It separates data based on features is akin to that of a tree model. With each iteration of the model fitting process, additional trees are incorporated to address and correct any prediction errors made by previous models. Each sample is allocated to a cluster of leaves in a tree that indicates a numerical weight based on the values of its input variables.

In order to enhance the performance of the XGBoost model, we utilized several hyperparameters. The hyperparameters for the model were configured as follows: the learning rate was set to 0.1, max depth to 17, n estimators to 200, subsample rate was set to 0.5, gamma value was set to 0, and the seed was set to 50. The hyperparameters were tuned to ensure the best possible performance of the XGBoost model, taking into account factors such as speed, accuracy, and flexibility.

2.2.3. Support vector machine

The SVM is a classification method that effectively reduces model complexity while accurately fitting the training data. The approach operates by identifying linear hyperplanes that maximize the distance between two sides or edges of several hyperplanes, thereby minimizing the likelihood of generalization errors [41]. Reducing the number of support vectors (data point closest to the hyperplane), helps to simplify the model and reduce overfitting. By doing so, the hyperplane becomes less dependent on these support vectors, which increases the potential for generalization.

To build an ideal separation plane from the starting data in a high-dimensional feature space, SVM employs an appropriate kernel function. In this study, we used a complexity constant of C = 5, to determine the tolerance for misclassification. A high value of C may lead to overfitting issues, while a moderate value may result in overgeneralization. Furthermore, in this study, we utilized a gamma value of 0.1 and a kernel of 'rbf'. The 'rbf' kernel function denotes the radial basis function, a widely employed technique in SVM to transform the data into a higher-dimensional space. The utilization of SVM enables the efficient identification of non-linear correlations among the features, thereby augmenting its precision and efficacy.

2.2.4. Random forest

The random forest (RF) offers a distinct advantage over other ML models by effectively reducing generalization error. This

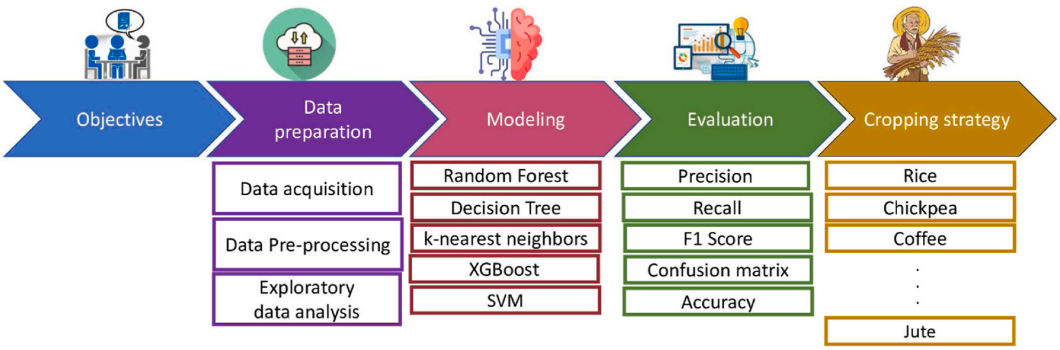

Fig. 1. Outline the comprehensive workflow for conducting analysis and recommending cropping strategies through the utilization of machine learning techniques.

Classifier, an ensemble learning method, is utilized to classify data by generating a forest consisting of a varying number of trees and subsequently averaging the predictions of each individual decision trees. Unlike basic DT algorithms, which are rule-based and rely solely on rules for predicting data sets, RF classifiers employ a different approach by randomly partitioning the functions instead of utilizing the Gini index or gaining weight to estimate the root-node. Each individual tree generates a prediction, and the ultimate outcome is determined by the class with the highest number of votes.

A criterion refers to a mathematical function that is utilized to evaluate the quality of a split. The entropy criterion is unique to trees, whereas the Gini criterion supports the Gini impurity. In this study, the RF model was tuned with the following parameters: a maximum depth of 6, maximum features of 5, minimum samples split of 4, random state of 0, and n estimator of 15.

2.2.5. K-nearest neighbors

The K-nearest neighbors' approach is a commonly used method for identifying the k number of training samples in a given training set that exhibits the highest degree of similarity to a target object [42]. Once identified, the approach assigns the dominant category to the target object by leveraging the category of the K training samples. When making a classification decision, the approach takes into account solely the category of the sample or samples that exhibit the closest proximity. The KNN approach is only applicable to a relatively limited set of neighboring samples for making category decisions, and it is not well-suited for handling high-dimensional data. In this study, the Minkowski distance formula was selected to compute the distance between neighboring data points. This particular formula has been demonstrated to exhibit superior accuracy compared to other distance metrics. The model was configured with a value of $k = 3$, indicating that the classification decisions were made based on the three nearest neighbors.

2.3. Training, testing and performance evaluation

The use of imbalanced data poses a major challenge in training and testing ML models as biased models tend to produce inaccurate predictions for the minority class. To address this issue, we employed an under-sampling strategy. In order to forecast the optimal

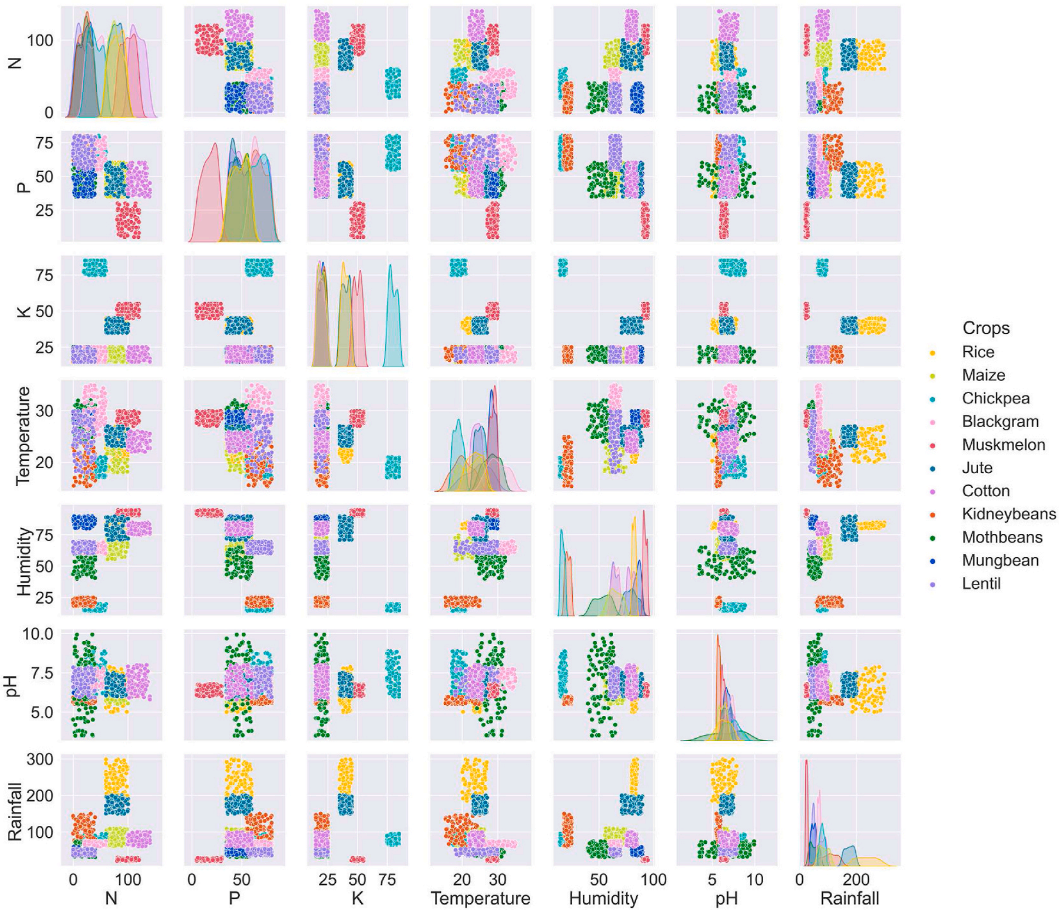

Fig. 2. Plot matrix showing the requirements of NPK, soil pH, temperature, humidity and rainfall for agricultural crops cultivating in different regions of India. Each cell in this plot displays the relationship between the variable corresponding to the row and column. The diagonal represents kernel density plots for each variable.

cropping strategies, machine learning algorithms were utilized. The training dataset consisted of 70 % of the available data, while the remaining 30 % was reserved for testing purposes. Effective model performance evaluation cannot be achieved solely by using accuracy. To assess the performance of the classifier models, additional model assessment metrics were generated in addition to accuracies, such as recall, specificity, precision, F1-score, and AUC based on [10,43]. AUC is the most efficient indicator for quantifying predictive power and was used to compare the accuracy of predictions.

3. Results

3.1. Requirements for nutrients and environmental factors of agriculture and horticulture crops

The plot matrixes show the requirement of NPK along with temperature, humidity, rainfall and soil pH for the cultivation of different agricultural (Fig. 2) and horticultural crops (Fig. 3). The temperature range for agricultural crops is generally narrow; however, certain crops such as black gram, moth bean, mung bean, and muskmelon require a relatively higher temperature. In regions with high humidity, crops such as muskmelon, mung bean, and rice are recommended due to their ability to grow well in such conditions. Conversely, in areas with relatively low humidity, crops such as chickpeas and kidney beans are more suitable for cultivation. Mung beans have the unique ability to thrive in a wide range of soil pH levels, unlike many other agricultural crops which typically require a pH range of 6–8. However, it should be noted that kidney beans exhibit a preference for slightly acidic soil.

Rice requires a medium level of NPK availability(Figs. 2 and 4(a-c)) in the soil with high rainfall (Fig. 4g) and natural p^H (Fig. 4f). The legume crops have a low Nitrogen requirement with a relatively high to medium-high Phosphorus requirement. Compared to the other legumes chickpea and black gram demonstrated relatively a higher demand for nitrogen (Fig. 4a). While comparing the phosphate requirement, it has been observed that among leguminosae crops, four of them namely black gram, chickpea, lentil, and kidney beans needs more phosphate (Fig. 4b). Conversely, muskmelon necessitates the lowest amount of phosphate with highest humidity (Fig. 4e). Legumes generally exhibit a lower demand for potassium with the exception of chickpeas, which requires a notably

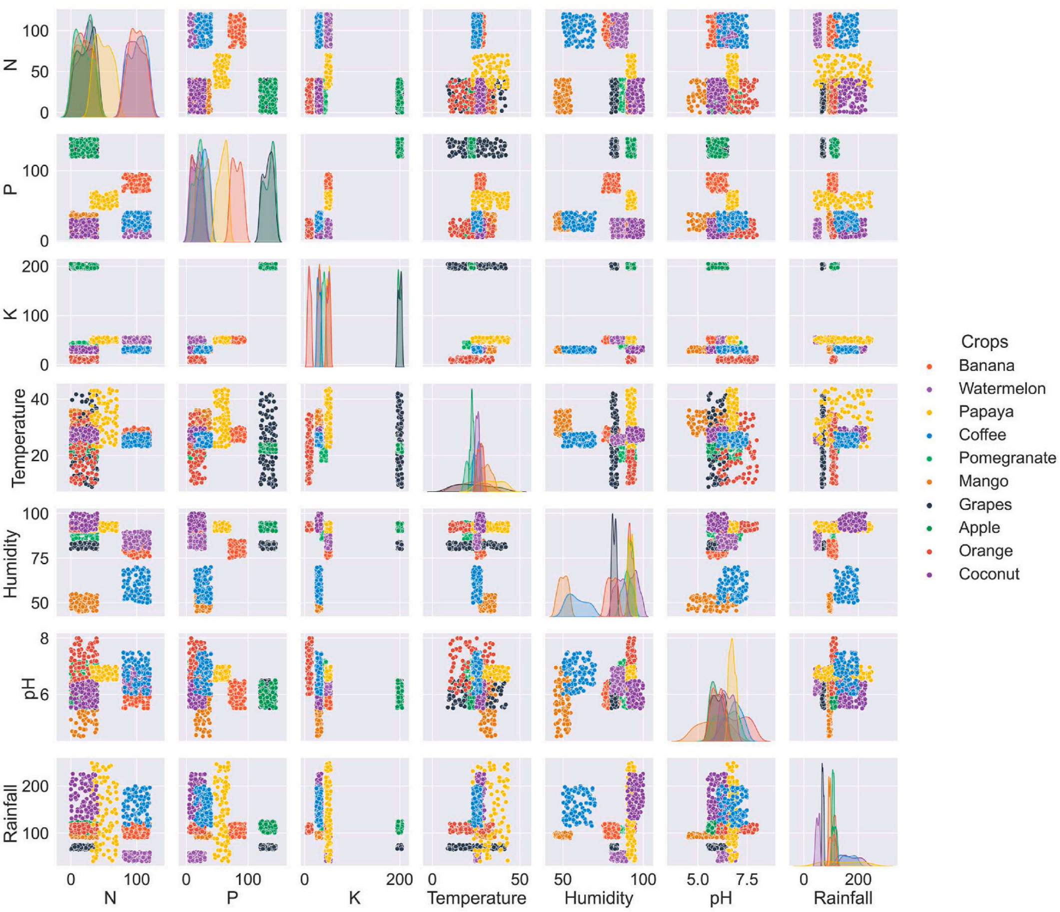

Fig. 3. Plot matrix showing the requirements of NPK, soil pH, temperature, humidity and rainfall for horticultural crops growing in different regions of India. Each cell in this plot displays the relationship between the variable corresponding to the row and column. The diagonal represents kernel density plots for each variable.

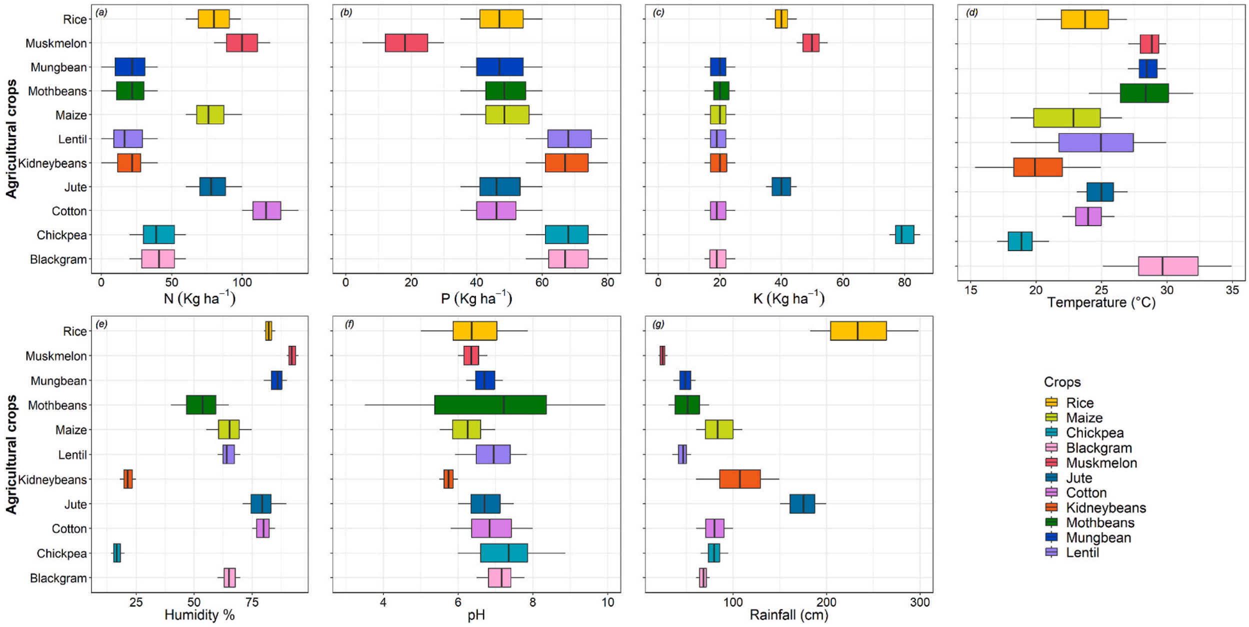

Fig. 4. NPK, soil pH, temperature, humidity, and rainfall requirements for 10 agricultural crops (a,b,c,d,e,f,g refers nitrogen, phosphate, potassium, temperature, humidity, pH, and rainfall for respective crops).

high amount of potassium (Fig. 4c). Cotton has been identified as the crop with the highest nitrogen and medium temperature requirement among all cultivated crops (Fig. 4a–d).

In the case of horticultural crops (Fig. 3), they showed a diversified climatic and edaphic requirement. Although their pH requirement was similar to the agricultural crops, they could grow in a wide range of temperatures. Horticultural crops like grapes and oranges exhibit a broad range of temperature tolerance, spanning from 10 °C to 40 °C. In contrast, the temperature requirements of most other horticultural crops fall within the range of 20 °C to 35 °C. Papaya is a versatile horticultural crop that can be grown in a wide range of regions, from those with low to high levels of rainfall (Fig. 5g). In contrast, most other crops typically require a moderate level of rainfall to grow successfully, with the exception of watermelon and grapes, which can still be cultivated in regions with as little

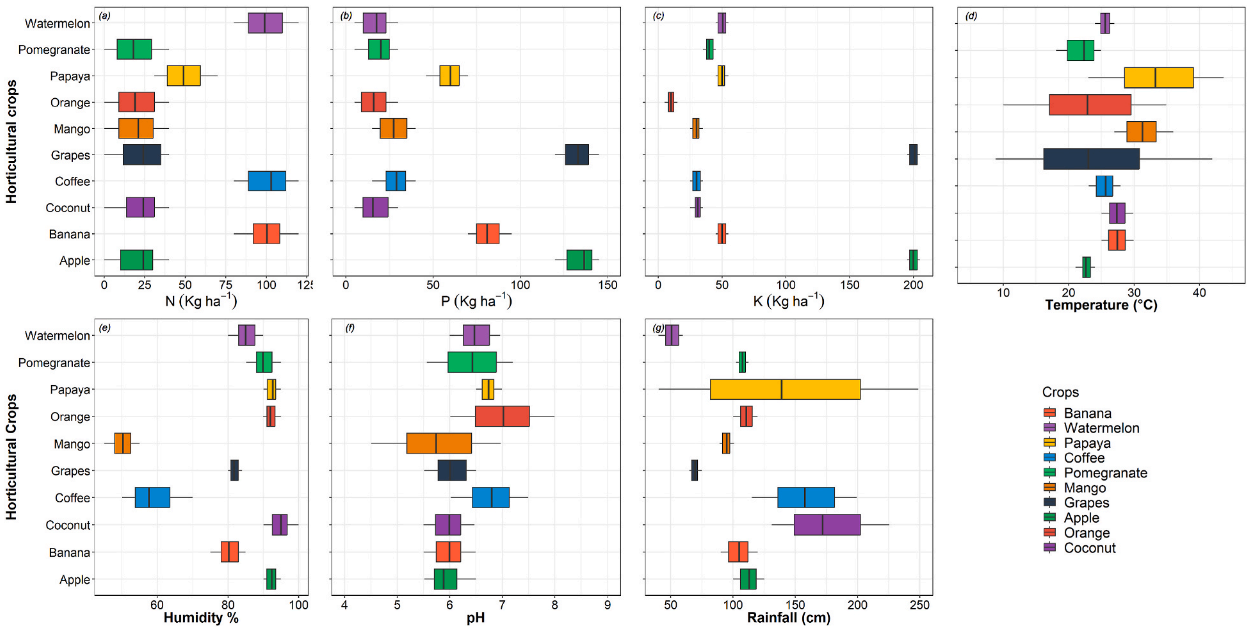

Fig. 5. NPK, soil pH, temperature, humidity, and rainfall requirements for 10 horticultural crops (a,b,c,d,e,f,g refers nitrogen, phosphate, potassium, temperature, humidity, pH, and rainfall for respective crops).

as 50 cm of rainfall and coconut requires a high amount of rainfall along with high humidity.

In the case of N, P and K conditions, apple exhibits the highest demand for nitrogen, phosphorus, and potassium (Fig. 5 a-c). The grapes showed a higher Potassium and Phosphorus requirement where it's nitrogen dependency was relatively lower than the others. The coconut also showed a lower requirement of NPK (Fig. 3). The majority of crops exhibit a preference for acidic soil, whereas mango and orange trees thrive in soil that is neutral to slightly alkaline.

3.2. Comparative analysis of ML classification performance

The study analyzed the performance trends of several machine learning algorithms, including random forest, decision tree, and support vector machine, in recommending the selection of agricultural and horticultural crops. The models were trained and fine-tuned according to the specific parameters detailed in the methodology section. Table-1 displays the predicted results for AC, HC, and mixed crop models, with regard to their test accuracy. The comparative performance of the models showed XGBoost demonstrated the highest test accuracy of 99.09 % and KNN demonstrated the lowest test accuracy of 94.45 % for agricultural crops. The model performances for horticultural crops exhibited almost equal levels of prediction accuracy, precision, and recall. In the context of using both AC and HC together in mixed crop modelling, it is noteworthy that the model yielded the highest accuracy of 98.5 %, which was comparatively lower than the accuracies achieved by using solely AC or HC in the models.

The evaluation of different models demonstrated that the AC-XGBoost and HC-XGBoost model exhibited superior performance compared to the AC-KNN and HC-KNN model, which obtained a comparatively lower score (Fig. 06a-b, Fig.S2.4-2.5). When evaluating a mixed crop model, the XGBoost model also outperformed the other models. Specifically, it achieved a perfect AUC score of 1.0 for 16 out of 21 crops, while the remaining 5 crops achieved high AUC scores ranging from 0.97 to 0.99 (Fig. 6c). Despite the fact that the majority of models attained elevated micro and macro average AUC scores of 0.99, the XGBoost model performed better in accurately predicting the classification of individual crops (Fig.S2.6).

3.3. Analyses of confusion matrix to visualize the recommendation response of the models

The evaluation of crop model performance is visualized through the analyses of confusion matrix. The matrix presented in Figs. 7–9 displays the true and predicted labels on the horizontal and vertical axes, respectively. The diagonal elements of the matrix represent the number of samples that were correctly identified. The AC model demonstrated superior recommendation performance for nearly all crops when utilizing XGBoost in comparison to other AC-ML models. However, XGBoost-AC model is not error free, it exhibited misclassification of black gram (5.8 %) as lentil and moth beans as shown in Fig. 7a. Similarly, the RF-AC model incorrectly suggested lentil (19.5 %) and moth beans (5 %) instead of black gram (Fig. 7b). This same model also exhibited misidentification of jute and lentil as rice and moth beans, respectively. There was a proclivity to erroneously recommend chickpeas and cotton as kidney beans, and maize respectively in all of the models except for XGBoost (Fig. 07, Fig.S2.1).

In the context of HC, all models exhibited misclassification of watermelon as mango and grapes as pomegranate (Fig. 8). The KNN algorithm exhibited misidentification rates of 6.25 % for coconut as orange, 4.54 % for grapes as pomegranate, and 3.22 % for pomegranate as orange, as shown in fig.S2.2c. The results indicate that the recommendations generated by XGBoost-HC (Fig. 8a) and RF-HC (Fig. 8b) outperformed those of the other methods, as illustrated in Fig. 8 and fig.S2.2. The mixed crop model that was combined exhibited a relatively higher rate of misclassification, as evidenced by Fig. 09a-b and Fig.S2.3. For instance, in the mixed crop-RF model, 5.5 % apple and 11 % coconut were misclassified as orange; lentil was misclassified as 10 % moth beans and 3 % mung beans, and 6 % papaya was misclassified as cotton (Fig. 09b).

4. Discussion

The selection of suitable plant species for cultivation is contingent upon the soil nutrient availability and the crop's demand for vital nutrients, such as nitrogen, phosphorus, and potassium (NPK), which are subject to the influence of diverse environmental factors, including soil pH, precipitation, temperature, and humidity. The experience of farmers commonly plays a pivotal role in selecting suitable crops for farming in a given site. However, the growing climate change and the deterioration of soil nutrient levels present significant challenges for farmers, leading to frequent instances of suboptimal decision-making. ML techniques are the most effective tools for providing crop farming recommendations, particularly with regards to the availability of soil nutrients and prevailing weather

Table 1
Comparative evaluation of the employed model's performance. The abbreviations AC, HC, and Co. are used to refer to these respective models of agricultural, horticultural, and mixed crop models.

Model	Accuracy (%)			Precision (%)			Recall (%)			F1- score (%)		
	AC	HC	Co.	AC	HC	Co.	AC	HC	Co.	AC	HC	Co.
RF	98.48	98.67	96.98	98.45	98.74	96.99	98.41	98.51	97.02	98.39	98.57	96.96
DT	94.55	98.00	87.62	94.30	98.05	88.32	94.26	98.09	88.52	94.15	97.99	88.07
SVM	96.67	98.2	95.71	96.72	98.40	95.72	96.55	98.19	95.68	96.36	98.23	95.60
XGBoost	99.09	99.3	98.51	99.05	98.09	97.33	98.75	98.90	97.99	98.11	97.91	97.49
KNN	94.45	97.67	92.24	94.03	97.71	92.06	94.31	97.56	93.18	94.31	97.54	92.00

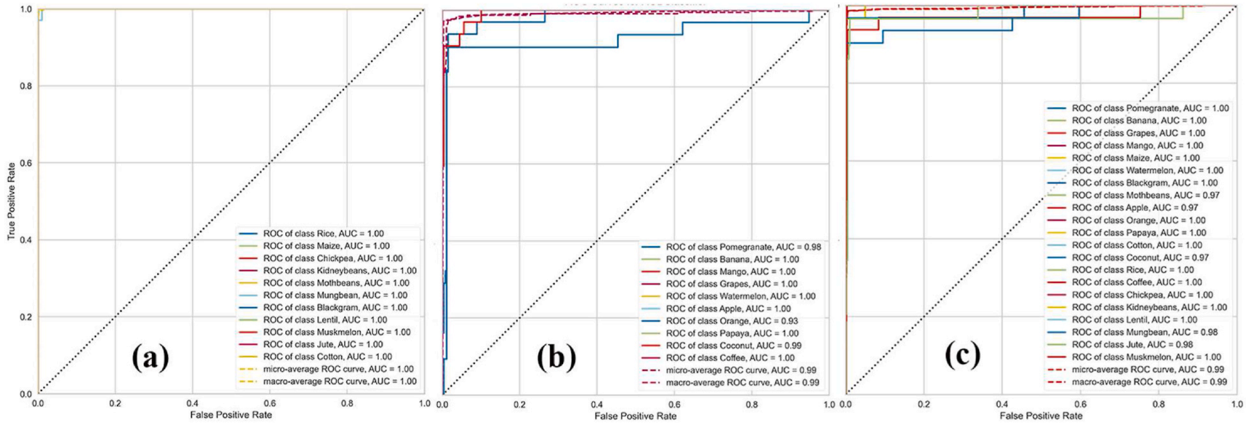

Fig. 6. ROC curves and the area under the curve (AUC) were obtained for agriculture crops (a), horticulture crops (b) and mixed crops (c) models using XGBoost algorithm.

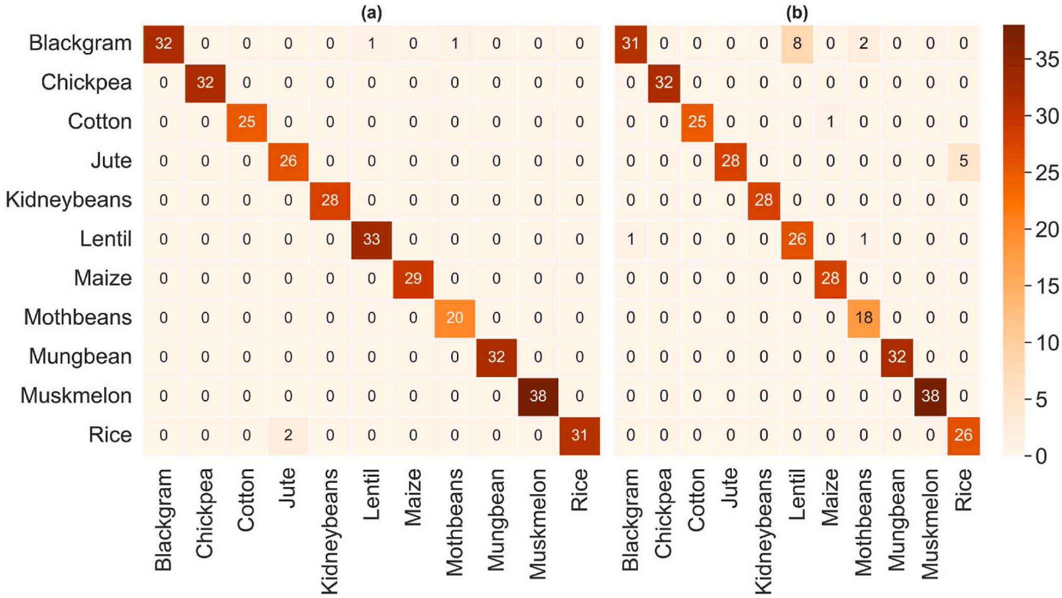

Fig. 7. Confusion matrices showing the best two machine learning models for agricultural crops where, a (left side) and b (right side) refers to results of XGBoost, and random forest algorithms respectively.

conditions. In the current study, we made an effort to evaluate the robustness of algorithms of ML for agricultural applications using authentic multi-trait field data collected from various locations in India, which was published in a public repository. The study illustrated that an ML-supported decision system, built on a non-mechanistic framework, can effectively suggest cropping across various environments. It should be noted that a multi-environmental dataset were used to achieve the outcome.

4.1. Performances of the models

The results indicate that XGBoost is a proficient model for predicting agricultural and horticultural crop categories, as demonstrated in fig. 06 and fig.S2.4-2.6. The KNN model showed lower accuracy in the classification of specific agricultural crops, such as Rice and Cotton, as evidenced by their respective AUC scores of 0.98 and 0.97, respectively. On the contrary, the XGBoost algorithm demonstrated superior performance in recommending horticultural crops, exhibiting a remarkable macro and micro average AUC of 99 % as depicted in fig. 06b. The XGBoost model demonstrated high accuracy in classifying various horticultural crops. Notably, for coconut, pomegranate, and orange, the models achieved impressive AUC scores of 0.99, 0.98, and 0.93, respectively while all other crop categories attained a perfect AUC score of 1.00. The accuracy levels are corroborated with the findings of another study [44] for crop recommendation and yield prediction where a recurrent neural network with red fox optimization techniques achieved ROC values ranging from 0.9 to 1.0. It is also supported by the study [45] where AUC values of 0.99, 0.98, 0.986, and 0.981 were obtained

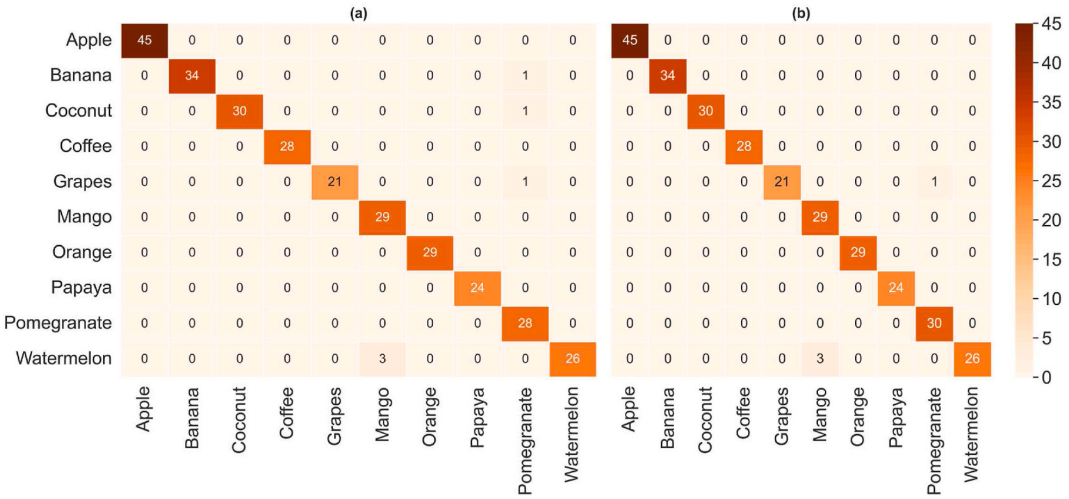

Fig. 8. Confusion matrices showing the best two machine learning models for horticulture crops where, a (left side) and b (right side) refers to the results of XGBoost, and random forest algorithms respectively.

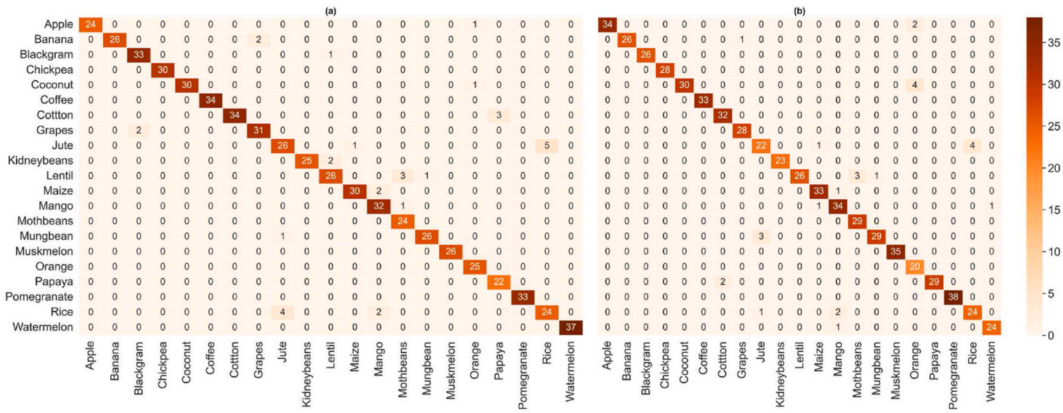

Fig. 9. Confusion matrices for mixed crops model of best two machine learning models, where, a,b refers to XGBoost, and random forest algorithms.

for XGBoost, RF, KNN, and DT, respectively, for crop classification and prediction based on soil nutrition. An ML-AI-enabled ensemble model [46] implemented for predicting agricultural yield reported AUC values of 0.72, 0.95, and 0.86 for KNN, SVM, and AdaBoost, respectively. This suggest the robustness of the models used the study for predicting crop suitability in different NPK regimes and climate zones in India.

4.2. Possible causes of miss-recommendation

The employed ML models exhibited variability in their performance. When predicting AC cropping, each model exhibited similar instances of incorrect recommendations (Fig. 07–09, and Fig.S2.1-2.3). Similarly, in some cases, rice was suggested as a substitute for jute because of their comparable NPK fertilizer and temperature requirements (Fig. 07) to attain the maximum crop yield. The confusion matrix revealed instances of misclassification between cotton and maize, likely due to their closeness in terms of rainfall, PK fertilizer, and temperature requirements.

The HC cropping also exhibited a similar phenomenon, as evidenced by the misclassification of orange as coconut due to their comparable NPK and humidity requirements (Fig. 5). The study revealed that the mixed crop model did not yield optimal accuracy in suggesting cropping strategies. When both agricultural and horticultural crop data were inputted into the model simultaneously, the confusion matrix revealed a chaotic scenario (Fig. 9 a-b, Fig.S2.3). The findings suggest that employing individual crop class modeling is more effective than combining different crop classes due to their comparable agronomic characteristics.

4.3. Implication of the study

Fig. 02–03 presents comprehensive nutrient (NPK) and climatic requirements for the cultivation of various agricultural and horticultural crops. These figures provide valuable insights into the specific elemental requirements of different crops, thereby offering a holistic understanding of the diverse agricultural and horticultural practices. In the realm of agricultural crops, rice and jute are known to have the highest demands for rainfall and humidity. This assertion is supported by Refs. [47,48]. Conversely, other crops tend to have lower to medium-low requirements for rainfall. Sridevi [49] reported an optimum RH of 70–80 %, which is consistent with our findings of 77 % RH (fig. 04e). Conversely, when the RH falls below 40 %, flowering is reduced. According to Ref. [50] legume crops exhibit a low demand for nitrogen, but require medium to high levels of phosphorus that support our findings. Additionally, their potassium requirements are diverse. It is widely recognized that crops exhibit optimal growth within specific ranges of temperature and pH. Horticultural crops exhibit diverse climatic, NPK fertilizer, and soil pH requirements, as illustrated in figs. 03 and 05a-g. The crops exhibited a wide range of temperature adaptability, with certain crops being suitable for high-temperature conditions while others were suitable for medium-high-temperature conditions. The nutrient requirements of different crops vary. For instance, grapes and apples have higher demands for potassium and phosphorus, whereas coconut necessitates lower levels of all three elements. Extensive model-based research is required to evaluate the demand for NPK doses by analyzing their interactions with environmental factors.

The notable implication of the study is the use of ML models and their robustness in identifying the right crops for a given site or the necessity of specific nutrients for growing a particular crop in a given environment. This is not the example of first study using ML, several studies have previously explored the use of ML models in the context of crop recommendation and yield prediction, each achieving varying levels of accuracy. For instance [44], SVM and DT models demonstrated accuracy rates of 91.73 % and 85.07 %, respectively. Another study [51], focused on smart farming and incorporated an intelligent insecticide and fertilizer recommendation system. This study achieved an accuracy of 78 % for SVM and 75 % for KNN models. Similarly, in Ref. [52] crop prediction was addressed by employing various feature selection techniques and machine learning classifiers. The study achieved noteworthy accuracy rates of 97 %, 90.6 %, and 88.83 % for RF, KNN, and SVM models, respectively. This was accomplished by selecting 8 attributes out of the total 15 using recursive feature elimination. Additionally [32], applied machine learning techniques to predict crops and reported accuracy performance ranging from 98.2 % to 88.5 % for the multilayer perceptron and JRip classifier. Another study [53] utilized random forest for sugarcane yield prediction, implementing forward feature selection and attaining an accuracy rate of 72 %. Furthermore [54], applied data mining techniques to recommend the planting of various crops in Bangladesh, utilizing both KNN and ANN models with impressive accuracy levels ranging from 90 % to 95 %. These findings are consistent with the results of other studies in the field, indicating that machine learning models hold promise for accurate crop recommendation and yield prediction. In this study, we have also observed higher prediction performance, aligning with the trends observed in these studies. The same dataset was previously used in two other studies [32,55] to make recommendations regarding appropriate crops based on various NPK and climatic variables. However, these studies [32,55] did not perform separate modeling of agricultural and horticultural crops. Additionally, they did not evaluate the comparative performance of ML algorithms through the use of confusion matrices, area under the curve (AUC), and ROC performance. All of these gaps have been addressed in the current study. In our comparative performance study of five different ML models, XGBoost demonstrated higher accuracy compared to the others. Specifically, for agricultural crops, XGBoost achieved an accuracy of 99.09 % and a macro average AUC of 1.0, as shown in Fig. 6. Similarly, for the horticulture trained model, XGBoost achieved an accuracy of 99.3 % and macro average AUC of 1.0, as depicted in fig. 06. In contrast to the 95.62 % accuracy reported by Thilakarathne [55], our study achieved a higher accuracy rate of 98.51 % for the XGBoost model (Table 1). Furthermore, our mixed model achieved a classification confidence of 99 %. The discrepancies observed between the current study and previous two studies [32,55] can be attributed to inadequate consideration given to the nutritional requirements, precipitation, temperature, RH and soil pH needs of each specific annual or perennial crop in those studies.

The implication of such modeling is for informed decision-making in practicing precision agriculture, as crop yield is influenced by a multitude of external factors. Attempting to comprehend the combined impact of these factors without the aid of modeling presents a formidable challenge. The agricultural productivity is significantly impeded by various biotic and abiotic stresses. It has been projected that environmental stress could lead to an annual economic loss of 0.3 %–0.8 % of the anticipated global gross domestic product in agriculture [56]. The selection of suitable crops presents an opportunity to tackle current challenges related to agricultural land use [57,58] and the global food shortage. The remarkable capability of ML is demonstrated in our study for rapidly recommending suitable crops based on varying geological settings and soil nutrient levels. The result of this study indicates ML-based recommendations for cultivating suitable crops using the public dataset of a given country can provide extensive services to end-users, ultimately resulting in optimal crop yields.

5. Conclusion

Precision agriculture has emerged as a transformative approach in the agricultural sector. The present study has made a significant contribution to this field by utilizing machine learning-based modeling to provide precise information on the suitability of agricultural or horticultural crops based on prevailing nutrient, climatic and soil pH variables. It is widely recognized that insufficient nutrient supply can have a detrimental impact on crop yield and may even lead to a decline in the soil's inherent capacity to sustain future crop cultivation. The agronomic requirements of agricultural crops vary from those of horticultural crops, particularly in terms of their nutrient needs (primarily NPK) and sensitivity to climatic factors. Hence, it is important to construct individual models for each class separately, as we have shown in our study. Based on the results of the study, it is recommended to use specifically trained models for individual crop classes in order to provide efficient, rapid, and more accurate cropping suggestions. The results of our study provide a useful reference for farmers in rural areas, as they can benefit from crop recommendations based on our findings, thus avoiding the need for trial-and-error farming. Moreover, the outcomes of our study can be utilized to design a user-friendly tool for a crop recommendation system that optimizes crop yield by taking into account the prevailing local environmental conditions. By offering more precise and accurate suggestions for crop selection and cultivation, it is possible to assist farmers in enhancing their productivity and efficiency, which can ultimately result in increased profitability and food security.

Despite the huge implications of the study, it should be noted that the authors in the present study used public dataset, more datasets across the globe could have resulted in more pragmatic findings from which the wider community could harvest the benefit. Therefore. It is recommended to set more experiments in the field in wider environmental conditions to generate quality data for modelling using various ML algorithms.

Funding

Funding institutions in the public, commercial, or nonprofit sectors did not award a specific grant for this research. However, Center for Research in Environment, iGen and Livelihood (CREGL) provide us substantial logistic support to accomplish the work.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Biplob Dey: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Jannatul Ferdous:** Writing – review & editing, Visualization, Formal analysis. **Romel Ahmed:** Writing – review & editing, Supervision, Project administration, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to express our gratitude to the Indian chamber of food and agriculture, the Indian agricultural extension department and anonymous farmers who support generating and gathering this robust dataset.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e25112.

References

[1] S.W. Wang, W.K. Lee, Y. Son, An assessment of climate change impacts and adaptation in South Asian agriculture, Int. J. Clim. Chang. Strateg. Manag. 9 (2017) 517–534, https://doi.org/10.1108/IJCCSM-05-2016-0069.
[2] K.K. Verma, X.P. Song, A. Joshi, D.D. Tian, V.D. Rajput, M. Singh, J. Arora, T. Minkina, Y.R. Li, Recent trends in nano-fertilizers for sustainable agriculture under climate change for global food security, Nanomaterials 12 (2022) 1–25, https://doi.org/10.3390/nano12010173.
[3] X. Liu, Y. Xu, S. Sun, X. Zhao, P. Wu, Y. Wang, What is the potential to improve food security by restructuring crops in Northwest China? J. Clean. Prod. 378 (2022) 134620 https://doi.org/10.1016/J.JCLEPRO.2022.134620.
[4] A. Bouguettaya, H. Zarzour, A. Kechida, A.M. Taberkit, Deep learning techniques to classify agricultural crops through UAV imagery: a review, Neural Comput. Appl. 34 (2022) 9511–9536, https://doi.org/10.1007/S00521-022-07104-9, 2022 3412.
[5] United Nations, World Population Prospects 2019, 2019. New York, https://population.un.org/wpp/Download/Standard/Population/.
[6] W.H. Meyers, N. Kalaitzandonakes, W.H. Meyers, N. Kalaitzandonakes, World population, Food Growth, and Food Security Challenges 15 (2015) 161–177, https://doi.org/10.1108/S1574-871520.
[7] Shoukat fiza, M. Awais, M. Khalid, Asad Sohail, The role of genetically-modified (GM) crops in food security, Life Sci. J. 19 (2022) 26–30, https://doi.org/10.7537/marslsj190222.04.
[8] G. Clarkson, P. Dorward, S. Poskitt, R.D. Stern, D. Nyirongo, K. Fara, J.M. Gathenya, C.G. Staub, A. Trotman, G. Nsengiyumva, F. Torgbor, D. Giraldo, Stimulating small-scale farmer innovation and adaptation with participatory integrated climate services for agriculture (PICSA): lessons from successful implementation in africa, Latin America, the caribbean and South Asia, Clim. Serv. 26 (2022) 100298, https://doi.org/10.1016/J.CLISER.2022.100298.
[9] R. Islam, R. Ahmed, B. Dey, S. Haque, S. Aktar, M.S. Bhuiyan, M.S. Arif, M.A. Habib Ador, M.M. Ul Haque, N. Saha, Salinity hazard drives the alteration of occupation, land use and ecosystem service in the coastal areas: evidence from the south-western coastal region of Bangladesh, Heliyon 9 (2023) e18512, https://doi.org/10.1016/j.heliyon.2023.e18512.
[10] B. Dey, M. Masum Ul Haque, R. Khatun, R. Ahmed, Comparative performance of four CNN-based deep learning variants in detecting Hispa pest, two fungal diseases, and NPK deficiency symptoms of rice (Oryza sativa), Comput. Electron. Agric. 202 (2022) 107340, https://doi.org/10.1016/j.compag.2022.107340.

[11] J.-H. Chen, J.-T. Wu, C. Young, The Combined Use of Chemical, Organic Fertilizers And/or Biofertilizer for Crop Growth and Soil Fertility, 2007, https://doi.org/10.30058/SE.200706.0001.
[12] M.A. Saleque, M.J. Abedin, N.I. Bhuiyan, S.K. Zaman, G.M. Panaullah, Long-term effects of inorganic and organic fertilizer sources on yield and nutrient accumulation of lowland rice, Field Crops Res. 86 (2004) 53–65, https://doi.org/10.1016/S0378-4290(03)00119-9.
[13] D. Batool, M. Shahbaz, H. Shahzad Asif, K. Shaukat, T.M. Alam, I.A. Hameed, Z. Ramzan, A. Waheed, H. Aljuaid, S. Luo, A hybrid approach to tea crop yield prediction using simulation models and machine learning, Plants 11 (2022) 1925, https://doi.org/10.3390/PLANTS11151925, 2022, Vol. 11, Page 1925.
[14] H. Burdett, C. Wellen, Statistical and machine learning methods for crop yield prediction in the context of precision agriculture, Precis. Agric. 23 (2022) 1553–1574, https://doi.org/10.1007/s11119-022-09897-0.
[15] S. Fei, M.A. Hassan, Y. Xiao, X. Su, Z. Chen, Q. Cheng, F. Duan, R. Chen, Y. Ma, UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat, Precis. Agric. (2022) 1–26, https://doi.org/10.1007/s11119-022-09938-8.
[16] D. Paudel, H. Boogaard, A. de Wit, M. van der Velde, M. Claverie, L. Nisini, S. Janssen, S. Osinga, I.N. Athanasiadis, Machine learning for regional crop yield forecasting in Europe, Field Crops Res. 276 (2022) 108377, https://doi.org/10.1016/J.FCR.2021.108377.
[17] A. Nihar, N.R. Patel, A. Danodia, Machine-learning-based regional yield forecasting for sugarcane crop in Uttar Pradesh, India, J. Indian Soc. Remote Sens. 50 (2022) 1519–1530, https://doi.org/10.1007/S12524-022-01549-0.
[18] R. Tufail, A. Ahmad, M.A. Javed, S.R. Ahmad, A machine learning approach for accurate crop type mapping using combined SAR and optical time series data, Adv. Sp. Res. 69 (2022) 331–346, https://doi.org/10.1016/J.ASR.2021.09.019.
[19] O.P. Duke, T. Alabi, N. Neeti, J. Adewopo, Comparison of UAV and SAR performance for Crop type classification using machine learning algorithms: a case study of humid forest ecology experimental research site of West Africa, Https://Doi.Org/10.1080/01431161.2022.2109444. 43 (2022) 4259–4286. https://doi.org/10.1080/01431161.2022.2109444.
[20] S. Nandhini, K. Ashokkumar, Machine learning technique for crop disease prediction through crop leaf image, Appl. Math. Inf. Sci. 16 (2022) 149–158, https://doi.org/10.18576/amis/160202.
[21] M. Sharif, M.A. Khan, Z. Iqbal, M.F. Azam, M.I.U. Lali, M.Y. Javed, Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection, Comput. Electron. Agric. 150 (2018) 220–234, https://doi.org/10.1016/J.COMPAG.2018.04.023.
[22] E.C. Too, L. Yujian, S. Njuki, L. Yingchun, A comparative study of fine-tuning deep learning models for plant disease identification, Comput. Electron. Agric. 161 (2019) 272–279, https://doi.org/10.1016/j.compag.2018.03.032.
[23] J. Chen, D. Zhang, A. Zeb, Y.A. Nanehkaran, Identification of rice plant diseases using lightweight attention networks, Expert Syst. Appl. 169 (2021), https://doi.org/10.1016/J.ESWA.2020.114514.
[24] H. Chu, C. Zhang, M. Wang, M. Gouda, X. Wei, Y. He, Y. Liu, Hyperspectral imaging with shallow convolutional neural networks (SCNN) predicts the early herbicide stress in wheat cultivars, J. Hazard Mater. 421 (2022) 126706, https://doi.org/10.1016/j.jhazmat.2021.126706.
[25] P. Bhatt, S. Sarangi, S. Pappula, Comparison of CNN models for application in crop health assessment with participatory sensing, GHTC 2017 - IEEE Glob. Humanit. Technol. Conf. Proc. (2017) 1–7, https://doi.org/10.1109/GHTC.2017.8239295, 2017-January.
[26] P.R. Sai Sankar, S.D.p.s Ramakrishna, M.M. Venkata Rakesh, P. Raja, V.T. Hoang, C. Szczepanski, Intelligent health assessment system for paddy crop using CNN, 2021 3rd, Int. Conf. Signal Process. Commun. ICPSC 2021 (2021) 382–387, https://doi.org/10.1109/ICSPC51351.2021.9451644.
[27] N. Kaur, Devendran, S. Verma, Kavita, N.Z. Jhanjhi, De-noising diseased plant leaf image, in: 2022 2nd Int. Conf. Comput. Inf. Technol., IEEE, Tabuk, Saudi Arabia, 2022, pp. 130–137, https://doi.org/10.1109/ICCIT52419.2022.9711604.
[28] Z. Doshi, S. Nadkarni, R. Agrawal, N. Shah, AgroConsultant: intelligent crop recommendation system using machine learning algorithms, Proc. - 2018 4th Int. Conf. Comput. Commun. Control Autom. ICCUBEA (2018), https://doi.org/10.1109/ICCUBEA.2018.8697349, 2018.
[29] V.C. Waikar, S.Y. Thorat, A.A. Ghute, P.P. Rajput, M.S. Shinde, Crop prediction based on soil classification using machine learning with classifier ensembling, Int. Res. J. Eng. Technol. 7 (2020) 4857–4861. www.irjet.net. (Accessed 23 December 2022).
[30] A. Dubois, F. Teytaud, S. Verel, Short term soil moisture forecasts for potato crop farming: a machine learning approach, Comput. Electron. Agric. 180 (2021) 105902, https://doi.org/10.1016/J.COMPAG.2020.105902.
[31] M.U. Ahmed, I. Hussain, Prediction of wheat production using machine learning algorithms in northern areas of Pakistan, Telecomm. Policy. 46 (2022) 102370, https://doi.org/10.1016/J.TELPOL.2022.102370.
[32] K. Bakthavatchalam, B. Karthik, V. Thiruvengadam, S. Muthal, D. Jose, K. Kotecha, V. Varadarajan, IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms, Technol. 10 (2022) 13, https://doi.org/10.3390/TECHNOLOGIES10010013, 2022, Vol. 10, Page 13.
[33] R.K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, A. Devare, Crop Recommendation System to Maximize Crop Yield Using Machine Learning Technique, Academia.Edu., 2017. https://www.academia.edu/download/55495855/IRJET-V4I12179.pdf. (Accessed 23 December 2022).
[34] N.H. Kulkarni, G.N. Srinivasan, B.M. Sagar, N.K. Cauvery, Improving crop productivity through A crop recommendation system using ensembling technique, Proc. 2018 3rd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. CSITSS (2018) 114–119, https://doi.org/10.1109/CSITSS.2018.8768790, 2018.
[35] D. Modi, A.V. Sutagundar, V. Yalavigi, A. Aravatagimath, Crop recommendation using machine learning algorithm, in: 2021 5th Int. Conf. Inf. Syst. Comput. Networks, 2021, https://doi.org/10.1109/ISCON52037.2021.9702392. ISCON 2021.
[36] P. Parameswari, N. Rajathi, K.J. Harshanaa, Machine learning approaches for crop recommendation, in: 2021 Int. Conf. Adv. Electr. Electron. Commun. Comput. Autom. ICAECA 2021, 2021, https://doi.org/10.1109/ICAECA52838.2021.9675480.
[37] 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 (2017) 950–953. www.irjet.net. (Accessed 23 December 2022).
[38] G. Suresh, D.A.S. Kumar, D.S. Lekashri, D.R. Manikandan, C.-O. Head, Efficient crop yield recommendation system using machine learning for digital farming, Int. J. Mod. Agric. 10 (2021) 906–914. http://modern-journals.com/index.php/ijma/article/view/688. (Accessed 23 December 2022).
[39] Kaggle. https://www.kaggle.com/, 2022. (Accessed 25 December 2022).
[40] ICFA, Indian Chamber of Food and Agriculture, 2022. https://www.icfa.org.in/. (Accessed 25 December 2022).
[41] Y. Hua, F. Li, S. Yang, Application of support vector machine model based on machine learning in art teaching, Wireless Commun. Mobile Comput. 2022 (2022), https://doi.org/10.1155/2022/7954589.
[42] B. Dey, K.A.M. Abir, R. Ahmed, M.A. Salam, M. Redowan, M.D. Miah, M.A. Iqbal, Monitoring groundwater potential dynamics of north-eastern Bengal Basin in Bangladesh using AHP-Machine learning approaches, Ecol. Indicat. 154 (2023) 110886, https://doi.org/10.1016/j.ecolind.2023.110886.
[43] B. Dey, R. Ahmed, J. Ferdous, M.M.U. Haque, R. Khatun, F.E. Hasan, S.N. Uddin, Automated plant species identification from the stomata images using deep neural network: a study of selected mangrove and freshwater swamp forest tree species of Bangladesh, Ecol. Inform. 75 (2023) 102128, https://doi.org/10.1016/J.ECOINF.2023.102128.
[44] P.S.S. Gopi, M. Karthikeyan, Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model, Multimed. Tool. Appl. (2023) 1–21, https://doi.org/10.1007/S11042-023-16113-2.
[45] T. Swathi, S. Sudha, Crop classification and prediction based on soil nutrition using machine learning methods, Int. J. Inf. Technol. 15 (2023) 2951–2960, https://doi.org/10.1007/S41870-023-01345-0.
[46] S.G. Kundu, A. Ghosh, A. Kundu, G. G P, a ml-ai enabled ensemble model for predicting agricultural yield, Cogent Food Agric. 8 (2022), https://doi.org/10.1080/23311932.2022.2085717.
[47] S.H. Ewaid, S.A. Abed, A. Chabuk, N. Al-Ansari, Water footprint of rice in Iraq, IOP Conf. Ser. Earth Environ. Sci. 722 (2021) 012008, https://doi.org/10.1088/1755-1315/722/1/012008.
[48] C.S. Murthy, M.K. Poddar, K.K. Choudhary, P. Srikanth, V. Pandey, S. Ramasubramanian, G.S. Kumar, Remote sensing based crop insurance for jute (Corchorus olitorius) crop in India, Remote Sens. Appl. Soc. Environ. 26 (2022) 100717, https://doi.org/10.1016/j.rsase.2022.100717.
[49] V. Sridevi, V. Chellamuthu, Impact of weather on rice – a review, Int. J. Appl. Res. 1 (2015) 825–831.

[50] H. Singh Jatav Sri Karan, J. Jena, S. Maitra, A. Hossain, B. Pramanick, H.I. Gitari, S. Praharaj, T. Shankar, J. Bharati Palai, A. Rathore, T. Kumar Mandal, H. Singh Jatav, Role of Legumes in Cropping System for Soil Ecosystem Improvement Improvement of Cropping System View Project Precision Agriculture View Project Role of Legumes in Cropping System for Soil Ecosystem Improvement, Nova Science Publishers, Inc., 2022.
[51] T. Thorat, B.K. Patle, S.K. Kashyap, Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming, Smart Agric. Technol. 3 (2023) 100114, https://doi.org/10.1016/J.ATECH.2022.100114.
[52] S.P. Raja, B. Sawicka, Z. Stamenkovic, G. Mariammal, Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers, IEEE Access 10 (2022) 23625–23641, https://doi.org/10.1109/ACCESS.2022.3154350.
[53] P. Charoen-Ung, P. Mittrapiyanuruk, Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning, Adv. Intell. Syst. Comput. 769 (2019) 33–42, https://doi.org/10.1007/978-3-319-93692-5_4.
[54] A.T.M.S. Ahamed, N.T. Mahmood, N. Hossain, M.T. Kabir, K. Das, F. Rahman, R.M. Rahman, Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh, in: 2015 IEEE/ACIS 16th Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distributed Comput. SNPD 2015 - Proc, 2015, https://doi.org/10.1109/SNPD.2015.7176185.
[55] N.N. Thilakarathne, M.S.A. Bakar, P.E. Abas, H. Yassin, A cloud enabled crop recommendation platform for machine learning-driven precision farming, Sensors 22 (2022), https://doi.org/10.3390/s22166299.
[56] M. Stevanović, A. Popp, H. Lotze-Campen, J.P. Dietrich, C. Müller, M. Bonsch, C. Schmitz, B.L. Bodirsky, F. Humpenöder, I. Weindl, The impact of high-end climate change on agricultural welfare, Sci. Adv. 2 (2016) 1–10, https://doi.org/10.1126/sciadv.1501452.
[57] L.K. Paine, P. Todd L, D.J. Undersander, K.C. Rineer, G.A. Bartelt, S.A. Temple, D.W. Sample, R.M. Klemme, Some ecological and socio-economic considerations for biomass energy crop production, Biomass Bioenergy 10 (1996) 231–242, https://doi.org/10.1016/0961-9534(95)00072-0.
[58] E.B. Barbier, Cash crops, food crops, and sustainability: the case of Indonesia, World Dev. 17 (1989) 879–895, https://doi.org/10.1016/0305-750X(89)90009-0.