National College of Ireland

Project Submission Sheet

Student Name: Ayush Jadhav, Anurag Singh, and Mohammad Amaan Shaikh……….

Student ID: 23178248, 23180013, and 23186925…………….……………..

Programme: MSC in Cloud Computing………..…**Year:** 2024-2025……………

Module: Cloud Machine Learning………………………………………………………………

Lecturer: Prof. Vikas Sahni………………………………………………………………………………

Submission Due Date: 22-07-24.…………………………………………………………………………………

Project Title: Crop Recommendation System………………………………………………………….

Word Count: 3174

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature: Ayush Jadhav, Anurag Singh, and Mohammad Amaan Shaikh

Date: 22-7-24

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Crop Recommendation System

Mohammad Amaan Shaikh
Student ID: 23186925
Email: x23186925@student.ncirl.ie

Anurag Singh
Student ID: 23180013
Email: x23180013@student.ncirl.ie

Ayush Jadhav
Student ID: 23178248
Email: x23178248@student.ncirl.ie

Project URL: https://colab.research.google.com/drive/1GRio91Zd0-wrcyySObsiWbs-tlpdYC-1#scrollTo=BLeKrTyUH_Ag
Dataset URL: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

Abstract—**There are still major constraints in agriculture associated with drastic changes in climate, the inability to restore the exhausted soil, and the lack of knowledge and practice regarding contemporary farming that puts annual production and farmer's income at risk. To address these issues, it is necessary to introduce methods based on advanced machine learning algorithms. In this study, five different ML models which are Support Vector Machine, K-Nearest Neighbors, Random Forest, Linear Discriminant Analysis, and XGBoost are tested by utilizing a dataset obtained from the Kaggle library that contains data on NPK, soil pH, temperature, rainfall, and humidity. Based on this, the paper involves suggesting 22 crops concerning Correlation Analysis with the usage of the MinMaxScalar technique, enhanced by methods including Dataset Sanity, Scaling, Confusion Matrix, Error Evaluation, and Data Visualization. More so, the study seeks to present practical solutions regarding the choice of crops to be grown, and the determinations of the crop nutrient, as well as serving to reduce the effects of climate in farming. Feature selection, error estimation, and results visualization are consistently performed to guarantee the model's efficacy. The study brings to light the fact that Random Forest possesses better capabilities as compared to the other classifications and obtains an accuracy level of 99.32 percent in predicting agricultural forecasts. This research helps to create crop recommendations depending on the environmental conditions and contributes to the further creation of an appealing AI cloud-based interface. An interface would be an opportunity to make fast decisions on the use of the fertilizers and choice of different crops in definite territories.**

Index Terms—**MinMaxScalar, NKP, Normalize, Correlation Analysis, Dataset Sanity, Scaling, Confusion Matrix, Error Evaluation, XGBBoost, Support Vector Machine, K-Nearest Neighbors, Random Forest and Linear Discriminant Analysis**

I. INTRODUCTION

Agriculture plays a crucial role in many economies and societies. Nonetheless, there are numerous challenges that farmers experience in crop production that significantly contribute to poor productivity, especially due to unfavorable weather and irregular rainfall. Other climatic problems result from poor fertility of the soil due to over-fertilization, and little access to information on modern farming techniques thus resulting in low production.

Crop production is affected by biotic and abiotic factors. Biotic factors include factors related to living organisms that are pathogenic to crops, these are bacteria, fungi, viruses, pests, and parasites which harm crops by way of predation and parasitism. Abiotic factors include soil type, atmospheric conditions, topography, water chemistry and sulfur oxides, nitrogen oxides, and heavy metals that have an impact on plant growth and yield. Another important aspect is how mechanical vibrations, radiation, and climate influence the productivity of agriculture. Suddenly weather changes also have adverse effects on the farmers and related activities all over the world, which in turn hampers the production of food crops, and other crops, and thus, economic imbalance.

Meeting the above challenges requires proper estimation of crop production for purposes of planning and resource allocation. Over the years, diverse forecast models to predict crop yield have been developed and are described in the literature. Myers et al. (Patel & Patel 2020) and Muriithi (Chlingaryan, Sukkarieh & Whelan 2018) explains how the use of subsequent analytical and mathematical procedures enhances crop prediction, which contributes to the creation of new products in the agricultural industry. In the paper of Muriithi (Chlingaryan et al. 2018), the author states that quantification of qualitative phenomenon produces more information about the situation, thus enhancing the accuracy of information and dependability of decisions to be made.

The main purpose of the present work is to give recommendations for choosing the crops to be grown and estimating the amounts of nutrients to use in combating the consequences of climate change on agricultural yields. To increase the precision of crop recommendation this research implemented data preprocessing comprised of data cleaning that involved removing duplicates, null values, and outliers, normalization by MinMaxScaler, and encoding categorical data. For data splitting Holdout method was used by following these steps: The features and targets of the data were divided into 80 % of the data used for training the model and 20 % was used to test the results. The random state used was 40 to make the results reproducible. The model selection was proposed based on the different core algorithms used by these models for predicting and accuracy levels. Programs such as scikit-learn which is in Python, numpy, and pandas were used in these processes while data visualization was done by Matplotlib and Plotly. The model's evaluation was done using a confusion matrix and error prediction. This comprehensive approach gives confidence that the results of this research shall be Reusable, Reliable, and Consistent for the Model training and Model evaluation processes.

II. Related Work

THis research notes that several prior papers concentrate on different elements of crop yield prediction and outline the potential use of machine learning in precision farming. (Dey, Ferdous & Ahmed 2024) Ali et al. (2024) mentioned the use of AI in precision agriculture and stressed the incorporation of machine learning algorithms to boost crop management and yield prediction. Their work supports the benefits of AI but further explains that AI is still a hard and expensive area to implement as another disadvantage.

However, there is another article with a more biological focus published in the Frontiers in Plant Science in 2023 (Hasan n.d.), which deals with plant stress reactions and their impact on yield. The study is strong in covering various topics related to plant physiology; however, the authors do not demonstrate how the presented artificial intelligence algorithms can be employed in practical machine learning use cases to build accurate models that can be utilized in agriculture.

Some articles discussed in IEEE relate to the use of machine learning in agriculture or in farming. For example, the paper by IEEE (2023) (Indira, Sobhana, Swaroop & Phani Kumar 2022) focuses on the application of CNN networks in crop diseases as well. Compared with categorical data, the method is very efficient in analysing image data; however, it requires a large amount of labeled data and high computational resources. Likewise, the study by Priyadarshini et al.(Vuyyala, Kona, Pusuluri, Variganji & Nenavathu 2023) on IoT-based smart agriculture systems brings insights on actual time surveillance and data gathering but fails to cover the forecasting part adequately, according to the IEEE norms.

The paper published in the IJCRT in the year 2024 (Ganesh Khillare n.d.) presents a state-of-the-art review on the application of ML for CY prediction with an objective analysis of the strengths and weaknesses of different approaches. This work has a positive element in that it provides a general source of methods such as regression, decision trees, and even neural networks. However, it frequently does not go into as much detail in critiquing these methods' translatable real-world considerations and applications. The CAB International Digital Library (2024) (Library n.d.) points to data quality preprocessing. One of the research areas attracting much attention in big data analysis is data preparation though this present work gives much attention to it while other works do not. While detailed information on structural adaptation and software revolution, its information on topics related to sophisticated forms of machine learning is scarce, which could have been useful for getting more extensive information. Past use of datasets, including those in the studies retrieved from Google Scholar, shows the employment of conventional LM models and the problems associated with large datasets. These works indicate that though the current approaches work to give basic predictions, they are extremely limited in both precision and versatility in terms of species and location.

In this project, replicating is expected to lend itself well to the approach as it directly reuses existing methods to set baseline performance. Thus, future work will focus on enriching these models by improving preprocessing and applying real-time data feed to predict more accurately and construct more robust models. This approach builds upon the assets of related works while fixing their weaknesses, including streamlining of data handling procedures and flexibility of the model.

III. Design and Methodology

This section of our report provides a detailed overview of data used and Methodologies used to implement the Machine learning model for "Crop Classification with Recommendation System". As shown in the fig 1, this model aims to make use of features and parameters available in the dataset to assist farmers in selecting suitable crops. In this section, a detailed explanation of the process of data preprocessing and the engineering of a feature set, data visualizations, scoring and evaluation, testing, training, and prediction are presented.

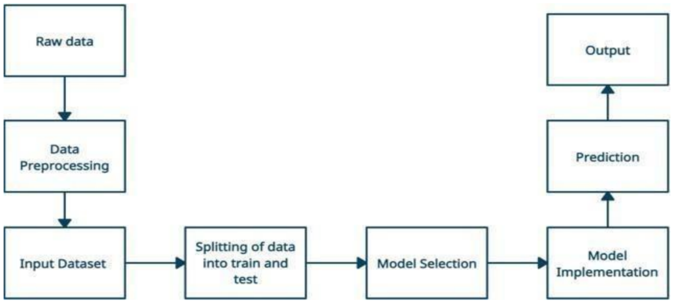


Fig. 1. Architecture Design

- **DataSet used:**
  The dataset used in this research project is sourced from Kaggle, which is referred to and collected over time by the government website "Indian Chamber of Food and Agriculture (ICFA)" (Ganesh Khillare n.d.). This dataset has a total of 2100 data entries where it involves 10 horticultural and 11 crops that are grown under different conditions as for the pH of the soil, NPK fertilization, and kind of climate such as humidity, temperature, and rain. This dataset is quite important because this dataset embraces a wide spectrum of crops and various geographical circumstances. It may therefore be applied to other regions of the world with similar characteristics in terms of the environment.

Data sets include:

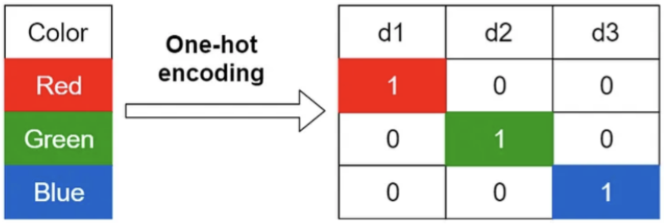
  - N - ratio of Nitrogen content in soil
  - P - ratio of Phosphorous content in the soil
  - K - ratio of Potassium content in soil
  - temperature - temperature in degrees Celsius
  - humidity - relative humidity in
  - ph - ph value of the soil
  - rainfall - rainfall in mm

- **Data Preprocessing:** Building an accurate Machine Learning model for crop recommendation requires important steps to be followed, which involve data cleaning, standardizing, and preparing raw data for training of models that will be used in the predictions. This section explains the methods to clean data, normalize data, encode categorical variables, and deal with outliers.
  - **Data Cleaning:** Data was cleaned to get rid of duplicate values, remove typos, etc. Duplicate values and null values were treated to make the data more accurate and consistent. Additionally, outliers were removed based on IQR (Interquartile Range) technique to check the spread of the middle 50(Lower bound) Q1 = numeric crop.quantile(0.25) (Upper bound) Q3 = numeric crop.quantile(0.75) Where Q1 represents or calculates the first quartile (25th percentile) for each numerical feature. And Q3 calculates the third quartile for each numerical feature. Furthermore, we calculated IQR using the formula: IQR = Q3 - Q1 Hence IQR helped us find the middle 50% of the data, lastly this data was filtered where any numerical parameter falling below (Q1 - 1.5 * IQR) or above (Q3 + 1.5 * IQR) is removed making our data outlier-free.
  - **Data Normalization and Standardization:** Normalization plays an important role in ensuring features are fitted and contributes to the model's learning process. The techniques that helped us achieve better accuracy after training were treating data with MinMaxScalar and then with StandardScalar.
  - **MinMaxScalar:** Before training and testing features were processed by MinMaxScaler bringing them to the range which is defined as [0, 1]. This scaler is particularly useful when the features to be scaled have different ranges and hence, are required to be normalized. The following formula defines the transformation that was done internally by the MinMaxScalar function of Python.

$$X' = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

    Where, X' is the normalized value, X is the original value, Xmin is the minimum value, Xmax is the maximum value.
  - **Encoding of Categorical Variables:** As many datasets have categorical variables, the machine learning algorithms work with numerical inputs. To get numerical from categorical data one-hot encoding was applied to retrieve number values. As shown in the fig 2, we have applied this technique to convert our column "label" from a categorical variable to n new binary variables. Where each category is represented as a vector, where the category to which it belongs is treated as "1" and all other positions will be marked "0", ensuring the model treats each label as a distinct numeric value.

    All the above preprocessing was implemented using Python's "scikit-learn", "numpy" and "pandas"



Color
Red
Green
Blue

d1	d2	d3
1	0	0
0	1	0
0	0	1

Fig. 2. Encoding Technique

- **Data Visualization:** When developing ML models and research visualization is essential. Fig 3 shows how environmental factors affects the production of crops in a particular region. For instance, it can be analyse that rice needs highest rainfall for its cultivation. It aids in the understanding of data patterns, model performance, and business insights. We have created a heatmap using Seaborn library to check the correlation of each feature, and plotted the distribution of each feature using Matplotlib, Plotly libraries, etc. The detailed understanding and discussion of visualization are discussed in the Research & Discussion session. Fig 4 shows correlation between independent attributes like Nitrogen, Phosphorus, Potassium, temperature, humidity, ph and rainfall. With the help of this matrix, which variable is more correlated to which variable or attribute, can be visualize.

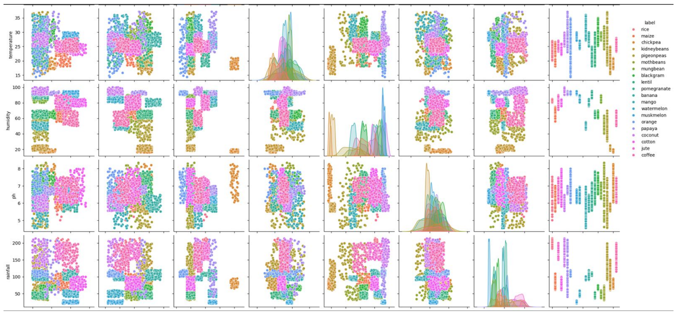


Fig. 3. Environmental factors affecting crop yeilds

- **Data Splitting:** Data splitting is an important step in the machine learning pipeline. It involves dividing the complete dataset into training and testing data. Proper data splitting will help divide the complete data well

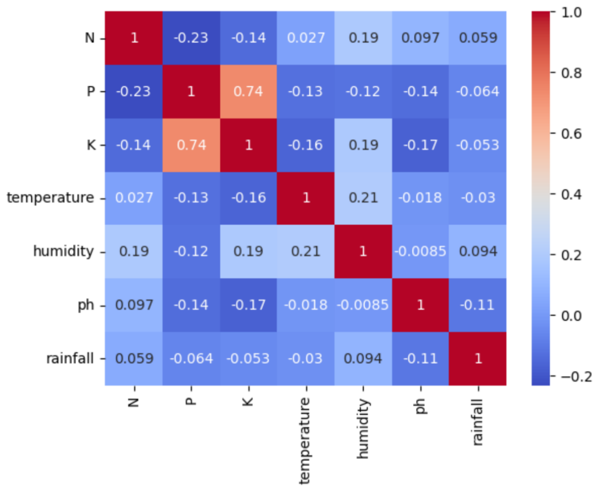


Fig. 4. Confusion matrix for Linear Discriminant Analysis

and help us get accuracy for test data. Based on which model is selected and can be used for the prediction. Training data: This data subset is used as data to train the machine learning model. Test data: It is used to evaluate model accuracy for the unseen data when made to learn on training data. There are various techniques for data splitting like the Holdout Method, K-Fold cross-validation, Stratified K-Fold Cross-Validation, Leave One Out Cross-Validation (LOOCV), Time Series Split, etc. Out of the above splitting methods we have used the Holdout Method which works on parameters like test size, random state, input variable, and output variable. (Anonymous 2024*a*)The holdout method is a simple and straightforward model evaluation technique that helps us to get reliable performance outcomes. In our research, we split data for 20% test data and 80% training set with a random state value of 40 making the results consistent.

- **Model selection:** We experimented with our model with 5 different machine learning algorithms: Support Vector Machine, K-Nearest Neighbors, Random Forest, Linear Discriminant Analysis, XGBoost. We chose these models as they showed the highest accuracy. For the final result and prediction, we are using random Forest as it showed the highest accuracy among all making it more compatible with our preprocessing.(Anonymous 2024*b*) The detailed evaluation, performance, and outcomes are discussed in the Result and Discussion section

IV. RESULT AND DISCUSSION:

To establish the performance of the various models an exhaustive assessment of different metrics was carried out in this study to predict the yields of crops. The used metrics As shown in the Table I include accuracy, balanced accuracy, precision, recall, and F1 score. These metrics give a comprehensive insight into the performance of the models developed in the current research by covering various sides of the aspects of predictability.

Accuracy computes the overall right classification rate in the whole range of data. It is expressed as the standard fraction of true positive and true negative rates divided by the total sample size.

Balanced Accuracy enhances proper detection accuracy based on class imbalances by calculating the average of the recall rate achieved for each of the classes.

Precision identifies the percentage of the true positive predictions out of all the positive predictions from the model. Recall (or sensitivity) means the percentage of correctly predicted positive among all real positive cases. **F1 Score** integrates precision and recall and allows you to receive a single coefficient that takes into account both indicators.

A. Implications and Comparative Analysis:

According to the analysis, it is evident that the Random Forest model provided the best simulation performance in all selected measures ensuring that it is the most stable and accurate model in predicting crop yields in this research. XGBoost also are equally good as Random Forest but it is slightly behind in the performance. These results point out that higher ensemble methods like Random Forest and XGBoost are reliable when dealing with the intricacy and volatility of the agricultural databases.

Support Vector Machine (SVM) and the K-Nearest Neighbors (KNN) were used but with less stability compared to the ensemble technique. In terms of appearance accuracy, especially in terms of distinguishing true positives, SVM was shown to be extremely precise. However, recall is slightly lower than precision, which means some positive cases were not identified. KNN showed equal performance with every evaluation criteria but was surpassed by the ensemble methods.

The lowest performance was noted for Linear Discriminant Analysis (LDA). This means that it may not be capable of Flagging all the positive samples and accounts for its relatively lower remember and F1-Score depending on the level of difficulty observed in handling the most complex and Imbalanced data sets.

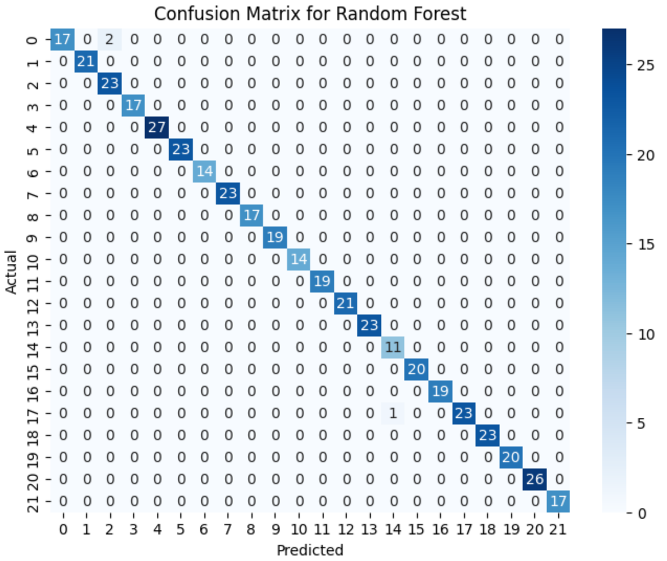
B. Error Analysis:

A detailed error analysis provides further insights into the models' performance:A detailed error analysis provides further insights into the models' performance:

Support Vector Machine (SVM): Precision at a high level, but recall a little lower – it shows that the positive cases are clearly defined, but some of them will be missed. Random Forest: All round very close to 1 indicating it is very robust and highly accurate. Linear Discriminant Analysis (LDA): Both recall and F1 scores used were low as they failed to locate several true positives and hence overemphasized the difference between the two. XGBoost: Still, it has 'high' performance on average, however slightly lower than Random Forest suggesting slightly lower efficiency in terms of classification.

Model	Accuracy	Balanced Accuracy	Precision	Recall	F1 Score
Support Vector Machine	0.9682	0.9695	0.9719	0.9682	0.9680
K-Nearest Neighbors	0.9591	0.9620	0.9654	0.9591	0.9590
Random Forest	0.9932	0.9933	0.9937	0.9932	0.9932
Linear Discriminant Analysis	0.9432	0.9448	0.9529	0.9432	0.9426
XGBoost	0.9864	0.9876	0.9869	0.9864	0.9863

TABLE I
SUMMARY OF MODEL PERFORMANCE METRICS



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	17	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	23	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17

Fig. 5. Confusion matrix for Random Forest

The confusion matrix visualizations focus on the models' capacity to rightly set the spots of the instances. This can be identify from the confusion matrix obtained in fig 5. Random Forest was again observed to have better performance, it had high values for true positives and true negatives which suggested the ideal application of the model in crop yield prediction.

C. Discussion:

The crops to be grown depend on aspects such as the nutrient content in the soils, plus other factors such as the pH level of the soil, rainfall, temperature, and humidity. Machine learning techniques serve as important instruments when it comes to recommending the crop farming best practices with the continually changing climate and deteriorating nutrient soil status.

All in all, it is noteworthy that great attention should be paid to the selection of the most reliable machine learning models for the advancement of precision agriculture. From this study, Random Forest and XGBoost were seen to give the best results as they provide accurate outcomes with relative ease and can indeed help farmers a lot in their decision-making process. The conclusions of the study support the further development of machine learning for agriculture and focus on possibilities for crop enhancement and growing productivity.

V. CONCLUSIONS

The project applies machine learning approaches to establish a crop recommendation system designed to assist farmers in deciding on which crop to grow depending on the characteristics of the soil and weather conditions like temperature, humidity, and rainfall. Five algorithms were adopted in building and evaluating the system including; Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), and XGBoost. Out of these methods, the Random Forest algorithm shows the best performance where its accuracy is 99. 32

Exploratory analysis was conducted using data visualization techniques to determine the key features related to the target variable. Metrics used in the performance and error assessment on the test set include the confusion matrix in addition to precision and the recall and F1 score.

The findings provided in the paper are rather unambiguous, and point to the fact that machine learning tends to be useful when it comes to crop recommendation systems. The system implemented in this project resolves problems associated with crop loss and increases the yield by recommending the most appropriate measure suitable to the existing environmental conditions for the improvement of crop control.

Therefore, the developed crop recommendation system can be considered a meaningful improvement in the approach to solving crop-related issues. Ideally, further research will compile new streams of data and use superior methods of machine learning to anticipate crop yields and suggest the best practices to achieve high production with low environmental repercussions.

REFERENCES

Anonymous (2024*a*), Data splitting methods for machine learning model training and testing, Technical Report TR-2024-001, Institute of Data Science, New York, NY.

Anonymous (2024*b*), Evaluating machine learning models for crop recommendation, *in* 'Proceedings of the International Conference on Agricultural Data Science', Agricultural Data Science Society, Chicago, IL, pp. 150–160.

Chlingaryan, A., Sukkarieh, S. & Whelan, B. (2018), 'Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review', *Computers and electronics in agriculture* **151**, 61–69.

Dey, B., Ferdous, J. & Ahmed, R. (2024), 'Machine learning based recommendation of agricultural and horticultural crop farming in india under the regime of npk, soil ph and three climatic variables', *Heliyon* **10**(3), e25112.

Ganesh Khillare, Omkar Kumbhakarna, S. M. P. (n.d.), 'ijcrt.org'. [Accessed 21-07-2024].

Hasan, M., M. M. A. U. M. P. A. M. I. K. S. M. S. . N. Y. . (n.d.), 'Frontiers — Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation — frontiersin.org'. [Accessed 21-07-2024].

Indira, D. N. V. S. L. S., Sobhana, M., Swaroop, A. H. L. & Phani Kumar, V. (2022), Krishi rakshan - a machine learning based new recommendation system to the farmer, *in* ‘2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)’, pp. 1798–1804.

Library, C. D. (n.d.). [Accessed 22-07-2024].

Patel, K. & Patel, H. B. (2020), ‘A state-of-the-art survey on recommendation system and prospective extensions’, *Computers and Electronics in Agriculture* **178**, 105779.

Vuyyala, V. R., Kona, M. S. R., Pusuluri, S. B., Variganji, S. & Nenavathu, B. (2023), Crop recommender system based on ensemble classifiers, *in* ‘2023 International Conference on Advancement in Computation Computer Technologies (InCACCT)’, pp. 68–73.

VI. Contribution Table

Name	**Responsibilities**
Anurag Singh	Design And Dataset
Ayush Jadhav	Experiment And Result
Mohammad Amaan Shaikh	Methodology
All individuals	Report