2023 International Conference on Integration of Computational Intelligent Systems (ICICIS)
Pune, India. Nov 1-4, 2023

Optimal Crop Rotation based on Soil Composition

Nithyashree Chandrashekar
MBA-Tech Computer Science
MPSTME NMIMS
Mumbai, India
nithya.shree01@nmims.edu.in

Rathin Nair
MBA-Tech Computer Science
MPSTME NMIMS
Mumbai, India
rathin.nair39@nmims.edu.in

Sanskriti Sharma
MBA-Tech Computer Science
MPSTME NMIMS
Mumbai, India
sanskriti.sharma75@nmims.edu.in

Tanmay Shende
MBA-Tech Computer Science
MPSTME NMIMS
Mumbai, India
tanmay.shende77@nmims.edu.in

Avani Bhuva
Computer Science
MPSTME NMIMS
Mumbai, India
avani.bhuva@nmims.edu

2023 International Conference on Integration of Computational Intelligent System (ICICIS) | 979-8-3503-1878-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICICIS56802.2023.10430231

***Abstract*— The agricultural sector has proven to be a critical sector for India since the early times. Even during recent years, we have seen that the agricultural sector provides employment to approximately 60 percent of the rural population. It has also contributed as significantly as 20 percent to the Indian GDP in the year 2020-2021. These reasons along with the fact that India has the largest quantity of agricultural land prove that the agriculture sector is one of the most essential sectors in India. Hence, we must ensure that the agricultural practices being employed, and the tools being used are up-to-date and that there is efficient use of arable land.**

In this paper, we reviewed various machine learning algorithms such as Artificial neural network, Random Forest, K-means clustering, K Nearest Neighbor (K-NN), Support vector machine, Neural Network, Decision Tree, Bayesian network etc. studied by various researchers to analyze the implementation with highest precision. In order to determine the most crucial parameters to take into account for the optimal decision, we also examined a variety of datasets and parameters considered by researchers for various algorithms. The aim is to determine the most suitable algorithm to help the farmers increase their crop yield, maintain soil nutrition, and gain more profits even during the off season. Hence, we test various algorithms on the dataset containing the most important parameters in order to determine the ideal crop.

Keywords*— *Machine learning, crop, yield, soil, nutrients, agriculture, crop rotation, fertilizer, KNN, Decision Tree, Random Forest, SVM, Neural networks, N, P, K.

I. Introduction

Agricultural sector is one of the most important sectors of occupation in India. It contributes greatly to employment, and to the GDP of the nation. Agriculture has also been a major contributor to the exports business. The agricultural sector also directly and indirectly supports various other sectors making it a critical part of the Indian economy. Hence, we must ensure that the agricultural sector performs as efficiently and effectively as possible with the resources given at hand. Various people have contributed to the research in this direction by introducing better irrigation systems, making artificial sunlight, etc. One such way of boosting the output of the agricultural sector is promoting Optimal Crop Rotation.

Crop Rotation is typically the practice of farming a pair of crops alternatively that help sustain the fertility of the soil and make sure that the land doesn't sit idle during off-season. It significantly improves the farmer's income and also makes sure that his land remains fertile for the upcoming crop. But due to lack of knowledge and understanding of the soil, many farmers make the wrong decisions regarding the pair of crops to be grown. This makes their land lose its fertility and become useless over time thus demanding fertilizers in a greater quantity. Hence it is of utmost importance that the crops selected by farmers are suitable to the soil to increase the yield and the profits of the farmer.

Crop selection for the soil can rely on diverse elements such as the temperature, pH level, levels of NPK and the amount of organic carbon content, etc. These factors must be taken into account according to their level of impact while making the decision of the ideal crop. There has been significant research on this front by many researchers. Our aim in this paper is to understand the work done by these researchers and to understand how they have selected the factors most critical for choosing the ideal crop. We also look at the various gaps and understand the scope of improvementin their research.

In the following sections we present our research on the current landscape of the agricultural situation in regard to crop rotation and measure of yield; and how the scientist and researchers have tackled the issues concerning the same. We also see the various algorithms and techniques used by researchers to solve the problem of crop selection and the accuracy for each algorithm/ technique used.

We see that though there are various methods and techniques that have been employed, the problem solution is still not structures. One reason for this is the variety of datasetsand sets of parameters being used.

The sections in the paper are as follows: The second section comprises of the literature survey done regarding the topic of crop rotation and crop yield prediction. The third section lists the methodology and implementation employed by us to determine the ideal crop based on all important environmental factors. The final section includes the discussion and the conclusion regarding the same.

II. Literature Survey

[1]tested various machine learning algorithms such as Recursive Feature Elimination, Classifier, k nearest neighbor, Decision Tree, Random Forest and so on to predict which crop should be grown in a particular soil under certain environmental conditions using features such as soil pH, carbon, nitrogen, rainfall and temperature. The Recursive Feature Elimination resulted in the highest accuracy.

[2]conducted a survey to understand how different analytical techniques such as Decision Tree, Random Forest, Bayesian network, Artificial neural network, Support vector machine, K-means clustering, expectation maximization algorithm, DBSCAN, Apriori algorithm can be used to forecast yield of a crop using various agrarian factors such as soil fertility, crop availability, climate, pests and crop diseases.

[3]aims to achieve two main goals: selecting the ideal crop to plant in the next crop rotation and adjusting the irrigation system of the land with attentive irrigation. Throughout the monitoring phase, data on the field's soil properties, including soil temperature, moisture, and air temperature, were gathered. A wireless sensor network (WSN) was developed to collect this information and deliver insight through sporadic cloud uploading. Long Short-Term Memory (LSTM) networks have been shown to be the efficient technique.

[4]advocated the use of predictive analysis to estimate agricultural output based on soil variety, soil ionic strength, amount of chemical used, area of crop seeded, and crop yield in tones. K-NN, Naive Bayes, Random Forest, and Bayes Net were among the several methods used, with Random Forest having the greatest accuracy.

[5]presented data mining methods like K-Nearest Neighbor and Naive Bayes to assess soil activity and categorize it depending on how nutrient-rich the soil would be for crop production, hence predicting crop output.

[6]developed a machine learning algorithm using neuro-fuzzy logic implemented in MATLAB by detecting the concentrations of urea, potassium, magnesium, pH, and nitrogen. This work assisted farmers in maximizing their yield by predicting suitable crops for cultivation as well as crop rotation.

[7]suggested using Multiobjective Evolutionary Algorithms (MOEA) to solve a multi-objective optimization problem while taking various areas and goals into consideration. The goals are to reduce overall input costs, increase soil nutrient retention, increase economic return, reduce economic risk, and encourage product variegation in consecutive growing seasons.

[8]proposed AgroConsultant, an intelligent system built on the mix of big data analytics and machine learning methods like Decision Tree, K Nearest Neighbor (K-NN), Random Forest, and Neural Network, was suggested by the study. This research seeks to help Indian farmers choose the best produce to cultivate based on the weather, location, soil traits, season, and temperature.

[9]developed the system using machine learning methods like KNN, Ensemble-based Models, Neural networks, Decision Trees for harvest forecast, and SVM for rainfall prediction with the intention of helping farmers. Based on soil characteristics and weather conditions, aiming at betterment of farmers. The system also offers information on the type and amount of fertilizers that are necessary, as well as the seedlings that must be planted. By doing this job, producers can boost their profit margins and prevent land pollution.

[10]recommended appropriate crops taking soil conditions into consideration using a variety of machine learning algorithms, including KNN, Gaussian support vector machine (SVM), and bragged tree. Data from their system was used to control the application of artificial fertilizers and prepare the area for the following planting.

[11]gathered data related to the crops' seed, including temperature, humidity, and moisture content to facilitate successful crop growth. Along with the software, a mobile application for Android is also being created. The application allows users to input parameters such as temperature, and their location is automatically determined, initiating the prediction process. This process is based on supervised learning, Naive Bayes, and Gaussian Naive Bayes, resulting in an accuracy rate of 97%.

[12]concentrated on the pragmatic implementation of several machine learning algorithms and their quantification. The findings indicate that Random Forest and XG boost are the most accurate. Furthermore, the research considers the variability in data from rainfall, season, area, and temperature datasets to identify a consistent pattern.

[13]identified KNN classifier and decision tree as potential machine learning algorithms. The aim was to incorporate additional attributes into the system and refine the results, leading to improved yields and the identification of various patterns for forecasts. This system is valuable in identifying which crops are suitable for specific regions. The parameters taken into consideration include soil type and pH level, humidity, temperature, wind, and rainfall.

[14]proposed a new approach called Crop Selection Method (CSM) to maximize the net yield rate of crops throughout the season. The development of the CSM involved considering various machine learning techniques, including ANN, SVM, KNN, Decision Tree Learning, Random Forest, Gradient Boosted Decision Tree, and Regularized Greedy Forest. The crop production rate is dependent on various factors, such as the region's geography, weather conditions, soil type, soil composition, and harvesting methods.

[15]devised Majority voting technique in order to determine what crop is best suited for precision agriculture which is the ultimate goal for the author as well as the farmers. The algorithms used for Majority Voting techniques are Random tree, CHAID, K-Nearest neighbor and Naïve bayes, algorithms to accurately determine the class under which each case of soil condition falls. Their efforts will aid farmers in planting the appropriate seed depending on soil needs to boost national output.

[16]worked at quantifying and emphasizing the importance of Potassium in the agricultural yield of the spices in India. The authors conducted various tests in order to find out the requirements of various spices in terms of the NPK nutrients and the recommended level of these nutrients in order to successfully cultivate these spices. Along with this the authors studied the amount of Potassium, Nitrogen and Phosphorus present in organic and inorganic fertilizers. The results from their work showed potassium sulfate were better sources for pepper, and potassium chloride accounts for 98% of potassium applied for spice crops.

[17]used k- means clustering in order to classify the soils in various districts based on environmental factors, biotic

inputs, the area for cultivation and the individual yield of crops. The use of data mining methods to agricultural data is the primary area of research in order to predict crop yields for important cereal crops in important areas of Bangladesh.

[18]developed a model based on data generated for crop rotation in the region of Muzaffarnagar. The authors used the Stochastic Optimization Function to determine the most suitable crop rotation based on factors like Temperature, pH, NPK and EC, etc. The parameters were weighted to determine the most suitable crop rotation pair more accurately.

III. IMPLEMENTATION

A. Proposed Architecture

In order to solve the existing drawbacks, we propose a new architecture considering the difficulties faced in the previous research papers. We have considered the parameters that have been shown to have maximum impact on the fertility of soil and hence in determining the ideal crop. The collection of data is brought into a consistent state by transforming the qualities utilized in it into a certain range, preventing anomalies.

The data is standardized using normalization, and any missing values are deleted. As the dataset is formatted, redundancy is reduced, which also makes the data more effective for processing predictions. Furthermore, we have conducted various pre-processing in order to better understand the influence of these parameters on the soil. Based on the previous information we perform the processes of feature extraction and feature selection in order to optimize our dataset by dropping unnecessary parameters.

We then split the dataset into training and testing in the ratio 70:30. Next we train various models using the algorithms listed and the training dataset and quantify the results of the same for comparison. We select the best models based on previous comparisons and train our model to achieve higher accuracy. The trained model is then applied to the testing dataset to predict the ideal crop for any given values of the parameters. After evaluating the performance of the model, we then trained it to achieve further accuracy on the test data. Finally, we incorporate the values from real world data to give accurate and quick predictions that allow the users to make informed decisions regarding their land.

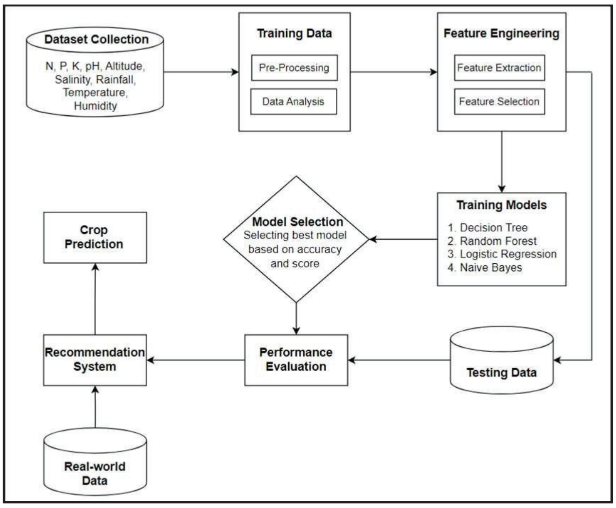

Fig. 1. Proposed Architecture

B. Data Collection and Description

The realization of our literature survey showed us that the reason for the existing inaccuracy of the current models was associated with the absence of consolidation of a few critical parameters. Further research revealed additional crucial parameters were soil salinity, Carbon content of soil and altitude. We referred the dataset [19] provided on Kaggle for the values of N, P, K, temperature, humidity, pH and rainfall. Additionally, we built the dataset for remaining parameters by augmenting the data available for the target crops.

The dataset encompasses a variety of crops, including grapes, muskmelon, pomegranate, watermelon, mango, orange, chickpea, lentil, mothbeans, mungbean, rice, cotton, kidney beans, apple, coffee, blackgram, banana, jute, pigeon peas, papaya, coconut, maize. Further information about our data is given in table 1.

TABLE I. DATA DESCRIPTION

Parameter	Description	Unit
N	Soil's Nitrogen concentration	mg/kg
P	Soil's Phosphorus concentration	mg/kg
K	Soil's Potassium concentration	mg/L
C	Organic Carbon concentration in the soil	%
pH	pH value of the soil	pH scale
Humidity	Relative atmosphericHumidity of region	%
Temperature	Temperature of the atmosphere in the region	Degree Celsius
Altitude	Altitude of the region	m
Rainfall	Rainfall in the region	mm

C. Training

After acquiring the dataset, we use the Logistic Regression, Random Forest, Decision tree and Naive Bayes Algorithms to train the model on the training data. This is done for selecting the model that optimally predicts the ideal crop. We then quantify the result of the trained models based on accuracy and select the best models.

- *Logistic Regression:* An instance of supervised learning is demonstrated here, aiming to predict the probability of a binary (yes/no) event. Logistic Regression is the chosen method for handling classification problems when dealing with a binary or categorical dependent variable. There are certain assumptions to consider while utilizing logistic regression, including the different types of logistic regression, independent variables, and the availability of training data.

- *Random Forest:* In the realm of classification and regression challenges, a frequently utilized supervised machine learning algorithm is the decision tree. This method constructs decision trees using diverse samples and leverages the majority vote for classification and the average for regression. The Random Forest Algorithm's primary advantage is its ability to handle both continuous and categorical variables in datasets, making it suitable for both regression and classification tasks. It performs better than other algorithms for these kinds of tasks.

- *Decision Tree:* Supervised learning technique that employs a tree structure to represent class labels and attribute representations. In this approach, the attribute of the record is compared to the root attribute, and based on the comparison's outcome, the algorithm moves to a new node. This process continues until a leaf node with a predicted class value is reached. Therefore, a decision tree model is highly effective in making predictions.
- *Naive bayes:* It is a supervised learning technique that creates a model of the class label distribution using a specific description of the target attribute. The naive Bayes method is based on applying Bayes' theorem and assumes independence between each pair of attributes. The Naive Bayes classifier reduces the complexity of the Bayesian classifier technique by assuming conditional dependency across the training set.

The results of the various algorithms are shown in the table.

TABLE II. TRAINING ACCURACY

Algorithm	Accuracy	Accuracy (%)
Decision Tree	0.8649350649350649	86%
Logistic Regression	0.8057142857142857	80%
Random Forest	0.8331168831168831	83%
Naive Bayes	1.0	100%

D. Testing and Results

Based on the accuracy of the algorithms on training data shown in the previous section we selected Decision Tree and Random Forest as the algorithms for our crop predictor model. We then moved on to the final phase of testing, where we applied both the trained models on the testing data created earlier. We then evaluated the performance of both models based on the confusion matrix as well as the performance metrics such as precision, recall, kappa, f1 score, accuracy. 22*30

The confusion matrix for the decision tree algorithm indicates that the diagonal elements are the correctly predicted samples. There are 563 such samples which are rightly predicted. There are a few samples, especially in classes 10, 18 and 20 that need to be trained more. The overall testing accuracy of the model that uses Decision Tree is 85.45%.

The confusion matrix for the Random Forest algorithm indicates that 529 samples are rightly predicted. There are a few samples, especially in classes 19 and 2, that need to be trained more. The overall testing accuracy of the model that uses Random Forest is 80.15%.

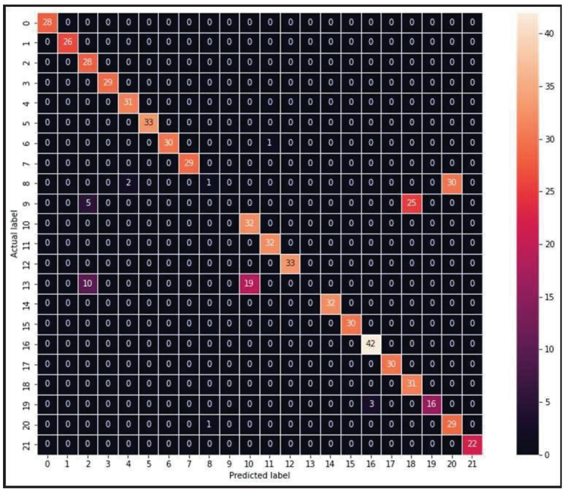

Fig. 2. Confusion Matrix for Decision Tree

```
DT3's Accuracy is:  85.45454545454545
              precision    recall  f1-score   support

       apple       1.00      1.00      1.00        28
      banana       1.00      1.00      1.00        26
   blackgram       0.65      1.00      0.79        28
    chickpea       1.00      1.00      1.00        29
     coconut       0.94      1.00      0.97        31
      coffee       1.00      1.00      1.00        33
      cotton       1.00      0.97      0.98        31
      grapes       1.00      1.00      1.00        29
        jute       0.50      0.03      0.06        33
 kidneybeans       0.00      0.00      0.00        30
      lentil       0.63      1.00      0.77        32
       maize       0.97      1.00      0.98        32
       mango       1.00      1.00      1.00        33
   mothbeans       0.00      0.00      0.00        29
    mungbean       1.00      1.00      1.00        32
   muskmelon       1.00      1.00      1.00        30
      orange       0.93      1.00      0.97        42
      papaya       1.00      1.00      1.00        30
  pigeonpeas       0.55      1.00      0.71        31
 pomegranate       1.00      0.84      0.91        19
        rice       0.49      0.97      0.65        30
  watermelon       1.00      1.00      1.00        22

    accuracy                           0.85       660
   macro avg       0.80      0.85      0.81       660
weighted avg       0.80      0.85      0.81       660
```

Fig. 3. Precision, recall, f1-score & support for Decision Tree

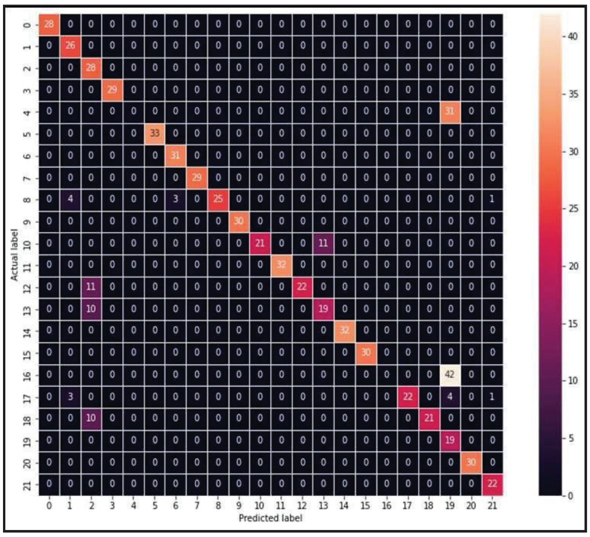

Fig. 4. Confusion Matrix for Random Forest

RF4's Accuracy is: 0.8015151515151515

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	28
banana	0.79	1.00	0.88	26
blackgram	0.47	1.00	0.64	28
chickpea	1.00	1.00	1.00	29
coconut	0.00	0.00	0.00	31
coffee	1.00	1.00	1.00	33
cotton	0.91	1.00	0.95	31
grapes	1.00	1.00	1.00	29
jute	1.00	0.76	0.86	33
kidneybeans	1.00	1.00	1.00	30
lentil	1.00	0.66	0.79	32
maize	1.00	1.00	1.00	32
mango	1.00	0.67	0.80	33
mothbeans	0.63	0.66	0.64	29
mungbean	1.00	1.00	1.00	32
muskmelon	1.00	1.00	1.00	30
orange	0.00	0.00	0.00	42
papaya	1.00	0.73	0.85	30
pigeonpeas	1.00	0.68	0.81	31
pomegranate	0.20	1.00	0.33	19
rice	1.00	1.00	1.00	30
watermelon	0.92	1.00	0.96	22
accuracy			0.80	660
macro avg	0.81	0.82	0.80	660
weighted avg	0.81	0.80	0.79	660

Fig. 5. Precision, recall, f1-score & support for Random Forest

IV. CONCLUSION

Machine learning models provide a considerable amount of scope to investigate the vast amounts of data, analyze the information obtained, and provide deeper understanding of the process Agriculture depends heavily on crop cultivation forecasts, with farmers eager to determine how much they may reasonably anticipate producing. Agricultural predictions in the past were based on farmers' fundamental knowledge of the crops that should be grown on plots of land. Considering the huge number of variable factors that can affect the growth of a particular crop and the impact of that crop on the soil, farmers might not be able to make very accurate predictions or take decisions based on the same. Using technologies like machine learning, we can get highly accurate results which can help farmers take informed decisions. Various tools are employed to create models that describe how various factors and activities interact. This makes it possible to predict potential responses to a particular scenario in the future.

The most often utilized algorithms are Decision Tree, K Nearest Neighbor KNN, Random Forest, Artificial Neural Networks, and Naive Bayes. Some studies have also employed feature selection to lessen the processing load on the algorithms and give users a more accurate answer in less time by removing redundant or unneeded components. Most of the publications that were examined considered variables like temperature, PH, NPK, moisture, rainfall, and so forth.

In our work we have successfully gathered data for various environmental parameters like N, P, K, C, pH, Humidity, Temperature, Altitude, Rainfall which play a vital role in determining the suitable crop. We also implemented 4 models before concluding that Decision tree model gives the best results for the given set of parameters.

V. FUTURE WORK

The implemented system demonstrates the ability to suggest suitable crops based on environmental factors. However, there are certain gaps that need to be addressed. Present model can be modified by incorporating other crucial attributes such as crop demand and supply, farm harvest prices, and market price as economic indicators. The multifaceted prediction will give farmers a ground level understanding of the best crops to cultivate by taking into account not only environmental and geographical elements but also economic aspects.

ACKNOWLEDGMENT

We would like to thank Prof. Avni Bhuva for her continued support and encouragement. We express our sincere appreciation for her generous contribution and guidance. We would also like to thank our college Mukesh Patel School of Technology Management and Engineering for enabling us to complete this research paper.

REFERENCES

[1] A. Suruliandi, G. Mariammal, and S. P. Raja, "Crop prediction based on soil and environmental characteristics using feature selection techniques," *Math Comput Model Dyn Syst*, vol. 27, no. 1, pp. 117–140, Jan. 2021, doi: 10.1080/13873954.2021.1882505.

[2] D. Elavarasan, D. R. Vincent, V. Sharma, A. Y. Zomaya, and K. Srinivasan, "Forecasting yield by integrating agrarian factors and machine learning models: A survey," *Comput Electron Agric*, vol. 155, pp. 257–282, 2018, doi: https://doi.org/10.1016/j.compag.2018.10.024.

[3] S. A. M. Varman, A. R. Baskaran, S. Aravindh, and E. Prabhu, "Deep Learning and IoT for Smart Agriculture Using WSN," in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2017, pp. 1–6. doi: 10.1109/ICCIC.2017.8524140.

[4] M. Chandraprabha and R. K. Dhanaraj, "Soil Based Prediction for Crop Yield using Predictive Analytics," in *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2021, pp. 265–270. doi: 10.1109/ICAC3N53548.2021.9725758.

[5] M. Paul, S. K. Vishwakarma, and A. Verma, "Analysis of Soil Behaviour and Prediction of Crop Yield Using Data Mining Approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 766–771. doi: 10.1109/CICN.2015.156.

[6] E. E. Vigneswaran and M. Selvaganesh, "Decision Support System for Crop Rotation Using Machine Learning," in *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, 2020, pp. 925–930. doi: 10.1109/ICISC47916.2020.9171120.

[7] R. Pavón, R. Brunelli, and C. von Lücken, "Determining Optimal Crop Rotations by Using Multiobjective Evolutionary Algorithms," in *Knowledge-Based and Intelligent Information and Engineering Systems*, J. D. Velásquez, S. A. Ríos, R. J. Howlett, and L. C. Jain, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 147–154.

[8] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018, pp. 1–6. doi: 10.1109/ICCUBEA.2018.8697349.

[9] M. N, "Crop Prediction using Machine Learning Approaches," *International Journal of Engineering Research and*, vol. V9, Aug. 2020, doi: 10.17577/IJERTV9IS080029.

[10] R. S. L. S. T. A. G. N. M.S. Roobini, "Crop suggestions using machine learning based on soil conditions," *nveo*, vol. 8, no. 5, 2021.

[11] M. Kalimuthu, P. Vaishnavi, and M. Kishore, "Crop Prediction using Machine Learning," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020, pp. 926–932. doi: 10.1109/ICSSIT48917.2020.9214190.

[12] A. Nigam, S. Garg, A. Agrawal, and P. Agrawal, "Crop Yield Prediction Using Machine Learning Algorithms," in *2019 Fifth International Conference on Image Information Processing (ICIIP)*, 2019, pp. 125–130. doi: 10.1109/ICIIP47207.2019.8985951.

[13] V. P. Prof. S. K. Pavan Patil, "Crop Prediction System using Machine Learning Algorithms," *International Research Journal of Engineering and Technology (IRJET)*, vol. 7, no. 2, 2020.

[14] R. Kumar, M. P. Singh, P. Kumar, and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," in *2015 International Conference on Smart*

Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015, pp. 138–145. doi: 10.1109/ICSTM.2015.7225403.

[15] A. R. Dasari, B. Dadore, and A. Watekar, “Crop Recommendation System to Maximize Crop Yield in Ramtek region using Machine Learning,” *Int J Sci Res Sci Technol*, pp. 485–489, Feb. 2019, doi: 10.32628/IJSRST196172.

[16] K. V. P. A. S. H. A.K. SADANANDAN, “Role of Potassium Nutrition in Improving Yield and Quality of Spice Crops in India,” *International Potash Institute*, 2001.

[17] A. T. M. S. Ahamed *et al.*, “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 International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 2015, pp. 1–6. doi: 10.1109/SNPD.2015.7176185.

[18] S. Aggarwal, R. Srinivas, H. Puppala, and J. Magner, “Integrated decision support for promoting crop rotation based sustainable agricultural management using geoinformatics and stochastic optimization,” *Comput Electron Agric*, vol. 200, p. 107213, 2022, doi: https://doi.org/10.1016/j.compag.2022.107213.

[19] A. Ingle, “Crop Recommendation Dataset,” Kaggle. Accessed: Feb. 28, 2023. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset/data