2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) | 979-8-3503-9737-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICACCS57279.2023.10112944

A Recommendation System for Crop Prediction Under Diverse Weather Conditions

ShunmugaLakshmi P
B.Tech Information Technology
National Engineering College
Kovilpatti, Tamil Nadu- 628503.
1915049@nec.edu.in

Ram Kishore V
B.Tech Information Technology
National Engineering College
Kovilpatti, Tamil Nadu- 628503.
1915403@nec.edu.in

Pillai Pratham Ramachandran
B.Tech Information Technology
National Engineering College
Kovilpatti, Tamil Nadu- 628503.
1915404@nec.edu.in

Santhi S
Dept. of Information Technology
National Engineering College
Kovilpatti, Tamil Nadu- 628503.
santhi_it@nec.edu.in

Kalaiselvi S
Dept. of Computer Science and Engineering
National Engineering College
Kovilpatti, Tamil Nadu- 628503.
sks@nec.edu.in

***Abstract*—In order to develop strategic initiatives in agricultural belonging for trading protocol and tripling countrymen earnings, early and accurate agricultural output estimation is pivotal in computable and commercial decision at the territory plane. Crop production forecasting, which is done to foresee a greater crop output, is the impressive difficulty in the agronomic division. To help farmers increase agricultural yield, this study has collected and evaluated data on N, P, K, temperature, humidity, ph, rainfall and soil_moisture. The superior values of this work are to recommend the crop to cultivate, suggest the appropriate fertilizers to use and predict the disease of a crop. A python environment is used to perform initial preprocessing on the data. KNN classification is used for developing the recommendation system. The link between the count, nutrients, temperature, humidity, pH, rainfall, and crop, are visualized using bar graphs and scatter plots. The tensor and torch were used to forecast crop disease and also to suggest appropriate fertilizer based on the data.The application was developed using Python Flask. This Classification was compared with other algorithms such as SVM, DT , RF and K-means**.

Keywords—Nitrogen, Phosphorous, Potassium, K-Nearest Neighbour, Support Vector Machine, Decision Tree, Random Forest and K-means clustering.

I. INTRODUCTION

Farmers and agriculture are harmed nationwide because they cannot produce adequate crops due to abrupt changes in the weather. Their inability to support their family forces them to take drastic action. The impacts of an increasing population taken together with the abnormal climatic conditions, soil abrasion, and an unstable geographical condition forces immediate plans to assure the timely and reliable crop growth and yield. Added to that, this causes the country's food resources to be scarcely available. Changes must be made to the circumstances faced by the farmers in our nation. Agriculture, which is the backbone of India's economy, has a significant impact. The agriculture industry employs the majority of the nation's labor force, which makes up more than 40%, and is directly or indirectly responsible for the livelihood of more than 50% of the nation. Every year, agriculture generates large amounts of data, so it is necessary to abandon the antiquated standard chart-based prediction methods in favor of a system that prioritizes and predicts outcomes using the available dataset.

The impact of the weather on crop productivity might be regarded as a top priority. Numerous studies have been done to determine how weather impacts agriculture; however, the majority of this research calls for very complicated information that is not readily available. This results in the acquisition of data through estimation, which can have a good or negative impact. Therefore, the methodology needs to be improved to account for the availability of data. Cultivators are able to forecast production template and crop impressions and then base crucial finding on them with the help of the graphical user interface .These suggest that the prediction of crop yields and crop protection are crucial to the global food supply.

The goal of a system is to produce a result on the basis of the functional impression and evident data. Since the result is based on the methodology and categories, recommender systems typically provide helpful advice. The Python flask is used to carry out the advice of crop. Data like humidity, rainfall, temperature, ph, soil_moisture, Nitrogen, Phosphorous, Potassium and air quality can be entered by the user, and the recommendation algorithm will then return the right crop to cultivate. The crop recommendation dataset is utilized to provide this recommendation. The crop-related data in the collection is gathered from open dataset sources. The KNN classifier classifies the data based on how similar it is. Tensors are used to divide an image's physical

characteristics into scalars and vectors. For image processing, a torch is used.

II.LITERATURE SURVEY

Ellavarasan et al, [1] this suggested work building a Deep Recurrent Q-Network model to anticipate crop yield, which is a RNN DL (Deep Learning) algorithm over the Q-Learning RL (Reinforcement Learning) method. The data parameters feed RNN''s progressively stacked layers. The Q-learning network creates a crop output forecast environment based on the information bounds. RNN output values are fed to Q-values using a linear layer. To predict crop yield, the reinforcement learning agent combines a threshold and a set of parametric characteristics. The net score for the actions taken is determined by minimizing error and maximizing forecast accuracy. With an accuracy of 93.7%, the suggested model accurately forecasts crop yield while outperforming other models and maintaining the original data distribution. In contrast Gupta et al, [2] In this study, the data is pre-processed in the Python environment before using the MapReduce architecture to continue processing , analyzing the substantial amount of information. Second, k-means was applied to MapReduce outputs and yields an average accuracy result for the data. After that, they examine the relationships between the two regions with cropName, rainfallInch, temperature of the region , soil of the region, and seed type of the region using bar graphs, scatter plots whereas this paper has the package of machine learning algorithms Rashid et al,[3] the proposed work is done by analyzing the enormous range of data and characterizing the output, Machine Learning architecture provide a clear understanding of the mechanism. Through the use of these technologies, models outlining a relationship between subject matter and actions are constructed. Additionally, using ML models, future responses in a specific circumstance can also be predicted. To acquire an overview, it's crucial to consider how frequently various ML algorithms are used and how well they perform. Linear Regression and Random Forest are the two most promising traditional ML systems. In addition to these techniques, some DL models, such as DNN(Dense Neural Network), CNN(Convolutional Neural Network), and LSTM, are used in the assessment of agricultural production on the other hand Alebele, Yeshanbele, et al, [4] their proposed work is to calculate the agricultural output from Sentinel-1A interferometric coherence and Sentinel-2A VIs, Gaussian Kernel Regression , Probabilistic Gaussian process regression—derivational from the Bayesian genesis of definitive GPR—are utilized.[5] This study looked into how interferometric coherence can add to the knowledge of optical data. The best score was recorded at the initial stage which is fed as input data for the regression models which is derived from the bitwise linear estimate of variable importance. It can be inferred that the GKR along with the sum of RDVI1, interferometric coherence metrics shall be of precise interest for a practical prediction of crop yield Likewise Devdatta et al, [7] propose a technique to forecast agricultural yield using historical data is proposed and put into practice. Using ML algorithms like SVM and RF on agricultural information,[6] this is accomplished and Fertilizer recommendations are made are appropriate for each of the crops. The research concentrated on developing a forecasting system that could be applied to agricultural output forecasting in the upcoming days. It gives a succinct description of how agricultural yield can be predicted using machine learning methods rather than regression algorithm, S. P Raja et al,[8] uses classification algorithms. The bagging strategy aids in properly predicting an appropriate crop while the modified recursive feature elimination functions chooses the most definite parameters. Several metrics, including accuracy, correctness, reminiscence, distinction, F1 score, the region under the curve, MAE(Mean Absolute Error) and log loss are used to assess the execution of the planned MRFE approach. The execution analysis justifies that MRFE technique outperforms other FS (feature selection) methods with a 95% ACC in contrast to Shetty et al, [9] The system was put into place so that people could learn more about farming and crops,[10] as well as how to harvest effectively. The study's primary focus is on agricultural datasets that were gathered from various portals that were located in certain districts of the state of Karnataka. well-organized datasets are presented. The accuracy of the crop yield forecast is achieved using the K-NN algorithm.

III.EXISTING SYSTEM

A method has been devised for predicting agricultural yields using an SVM classifier, DT, K-means and naive bayes. Comparing several techniques enables the selection of the methodology that is most appropriate for given forecast results in determining the best crop to grow by applying KNN classification to our final generated datasets. The existing algorithms are compared with KNN by accuracy. The crop is recommended with parameters such as temperature, rainfall, soil type. The number of crops that can be predicted is 120.

IV.PROPOSED SYSTEM

In the Proposed system, KNN Classification is used to forecast the crop to cultivate. Firstly, the collected raw data is pre-processed in a Python environment. Then this data is classified using KNN Classification. The python flask is used to develop a web-based application (http://127.0.0.1:5000/). When a user clicks an event, an API is called and user data is classified with the Actual data. The Technical Architecture is seen in Fig. 1.

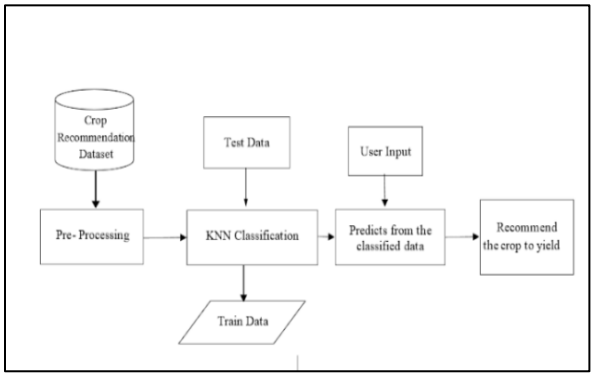

Fig. 1. Weather-Based Crop Prediction using KNN

A. Dataset

The crop-recommendation dataset is collected from open source (Dataset). The dataset includes temperature, rainfall, humidity, pH, N, P, and K.

B. Pre-Processing

The data is pre-processed with the python environment. Using StandardScaler of the Sklearn library the raw data are processed such that NA (Not Available) values in all column records are eliminated. The data before processing is seen in Fig. 2. The parameters with null data are replaced with the mean of that feature. The data after pre-processing is seen in Fig. 3[11]

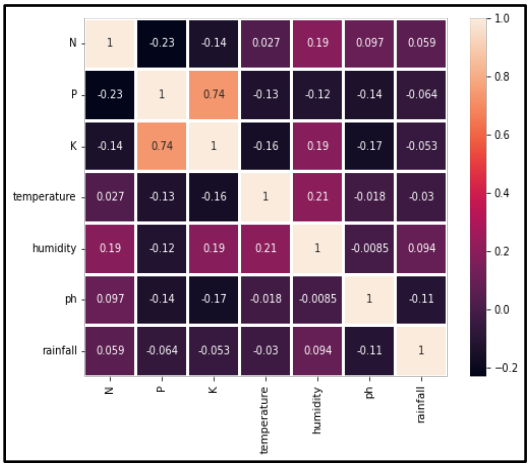

Fig. 2. Data before Pre-Processing.

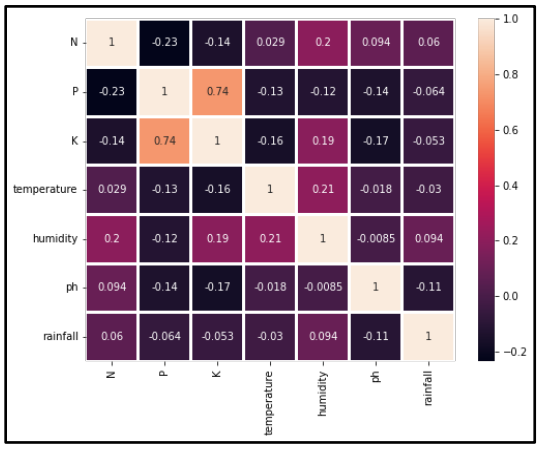

Fig. 3. Data after Pre-Processing.

C. Classification

The KNN algorithm classifies the data according to the similarity between the newly entered data and available data which in turn determines the similarity of this newly entered data and the available data. The newly entered data can be easily classified into a good suite category by KNN. Several alternative distance functions can be used to estimate the separation between class in a feature space, but the Euclidean distance function is the one that is most frequently applied[12]

When determining the separation between a and b, the

If x = (x1, x2) and y =, then the Euclidean metric is typically utilized.

When (y1, y2), the following distance is calculated:

$$dis(x,y) = y1 - x11 \qquad (1)$$

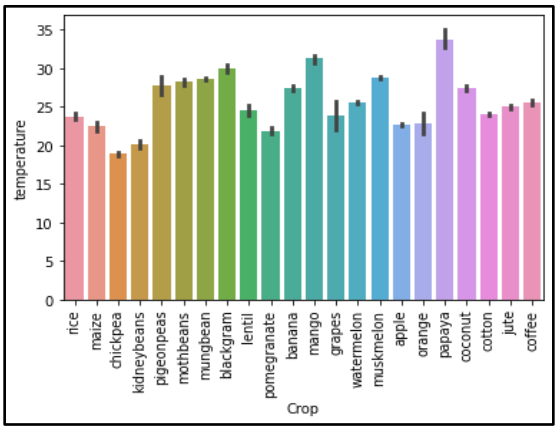

Fig. 4. Crops Vs Temperature

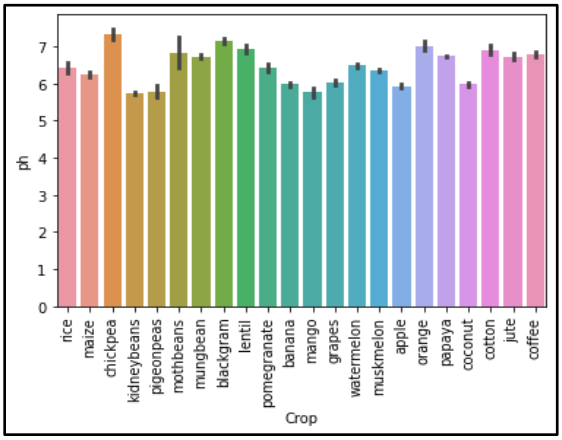

Fig. 5. Crops Vs pH

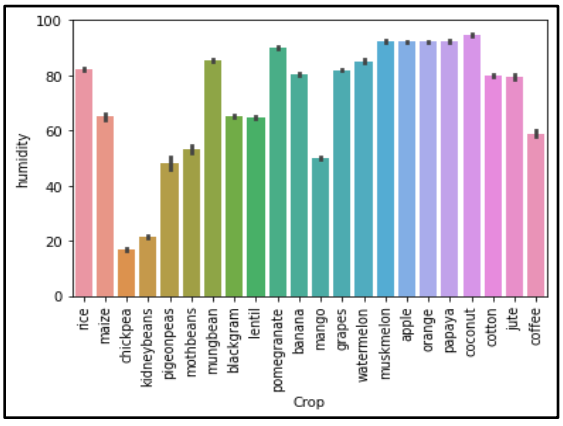

Fig. 6. Crops Vs humidity

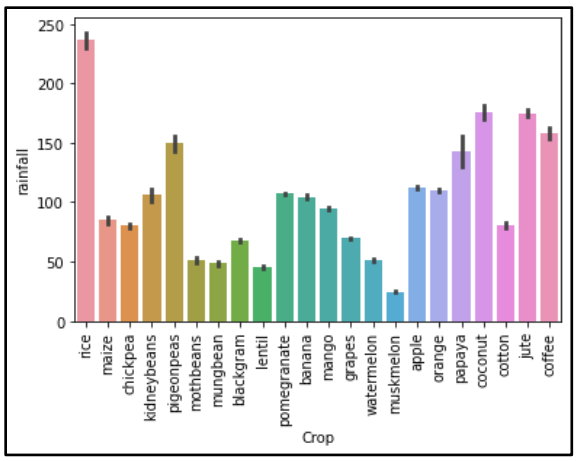

Fig. 7. Crops Vs rainfall

Each crop is compared with different weather conditions. It shows how each weather conditions influence in predicting the right crop. It is seen in Fig 4, 5, 6, 7. Accuracy gained is shown in Fig. 8.

Selecting the Feature and Target Variables:

```
x = df[['N', 'P','K','temperature', 'humidity', 'ph', 'rainfall']]
final = df['Crop']
```

Encoding target variable

```
y = pd.get_dummies(target)
y
```

Splitting data set into 25% and test dataset into 75%

x_train : (1650, 7)

x_test : (550, 7)

y_train : (1650, 22)

y_test : (550, 22)

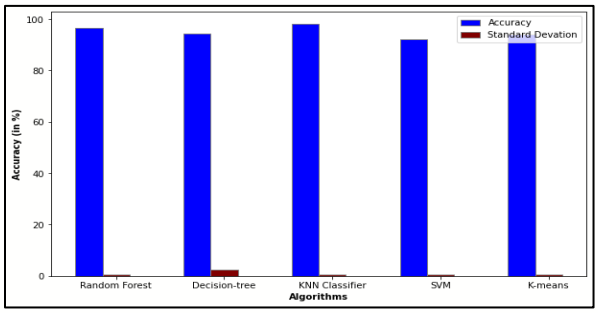

Fig. 8. Accuracy Gained

D. Python Crop Recommendation System

The Python recommendation system is built using python Flask. The user enters data such as temperature, rainfall, humidity, pH, country, state, and air and gets submitted. This post request is sent via an API. The response is a prediction. The crop recommendation system is seen in Fig 9. [13]

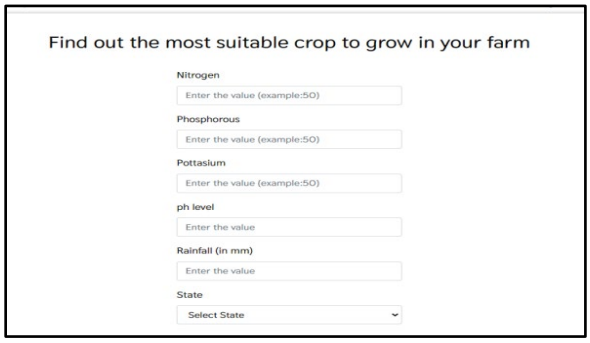

Fig. 9. Crop Recommendation Form

E. Python Fertilizer Suggestion

The fertilizer needed for the crop is suggested by the fertilizer dictionary data. This dictionary data has the fields such as crop, N, P, K, pH, and soil_moisture. The user enters a value for these fields and the post request is sent via an API, which suggests the nutrient deficiency or the nutrients needed for the particular crop. The user input for fertilizer recommendation system is seen in Fig 10.[14]

Fig. 10. Fertilizer Recommendation Form

F. Python Crop Disease Prediction

The user needs to upload a picture with any crop-diseased image and predict the disease of that particular crop. This is achieved using PyTorch and Tensor. The disease of the crop is predicted using this recommendation system is seen in Fig 11.

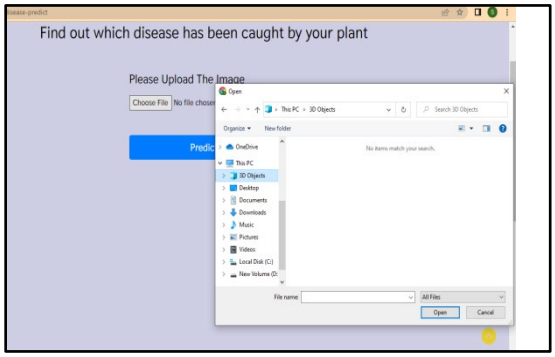

Fig. 11. Crop Disease Prediction Form

G. Website

A graphic user interface for the user is created for a simple website. Python 3, HTML 5.0, CSS3, and Bootstrap 4 were used to create the website's front end. using Flask 2.0.0 in a Virtual Environment. [15]

V. Results And Discussion

This work suggests a "Crop Recommendation System" and uses a K-Nearest Neighbor Classifier which gives valuable results. This recommendation system suggests the crop to yield, in addition, the recommendation system predict the crop disease and also suggest the fertilizer to use. The crop recommendation result is seen in Fig. 12. The fertilizer suggestion is seen in Fig. 13. The crop disease prediction is seen in Fig. 14. The maximum number of Crop predicted is 150. The number of records used is 2201 for predicting the right crop to yield. To predict the disease of the crop 36 classes were trained. For fertilizer suggestions 1507 records are used. The accuracy gained using the crop recommendation dataset is seen in Fig 8. The results for crop, fertilizer, crop disease is seen in Fig 12,13,14 respectively.

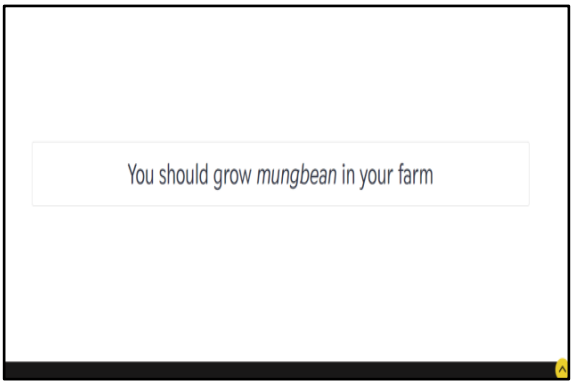

Fig. 12. Result of Crop Recommendation

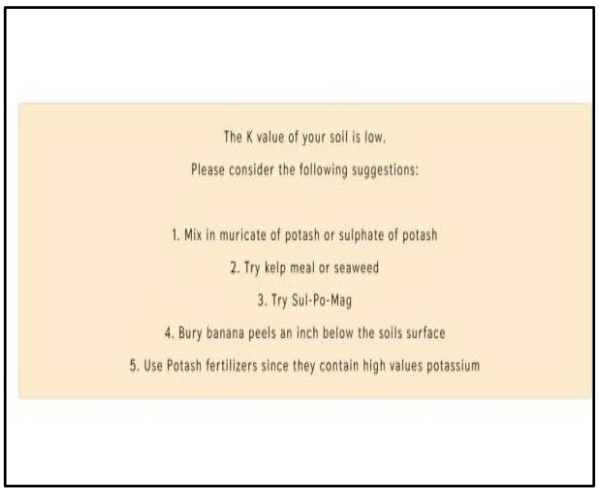

Fig. 13. Recommendation of a fertilizer

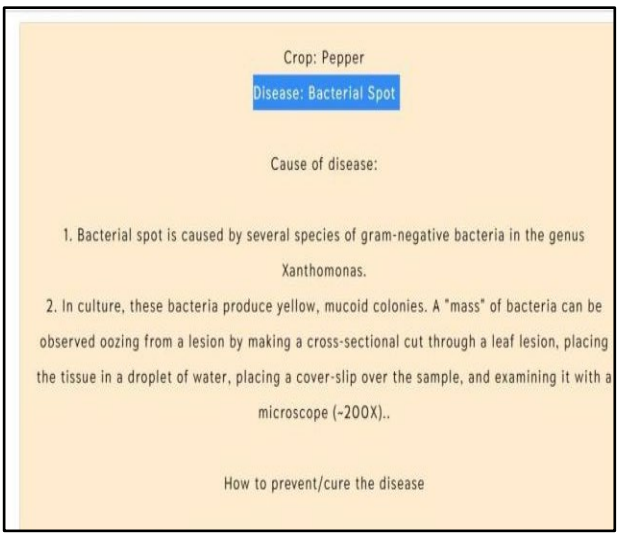

Fig. 14. Prediction of crop disease

VI. Conclusion

This work provides a "Crop Recommendation System" and uses a K-Nearest Neighbor Classifier which gives valuable results. The number of crops predicted using KNN is 150. The accuracy gained is greater than SVM, DT, naive Bayes and K-means. The web page design is easy for the end user to perform actions. The recommendation system with three major roles of prediction and suggestion is useful in increasing agricultural output. More records are used to provide the appropriate result. The data with NA and null values are pre-processed.

In the future, it is planned to combine the classification algorithm to achieve better results. The KNN is used with Multiple Linear Regression since in the previous studies' Multiple Linear regression also predicts maximum crops. Combining these two algorithms will ensure better accuracy i.e., a more number of crops predicted. The Dataset with even more features is planned to include. The input entered by the user is automated with the sensors. The inputs are derived from the Temperature Sensor, Humidity Sensor, Acid analyser etc.

References

[1] Elavarasan, Dhivya, and PM Durairaj Vincent. "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications." vol. 8, pp. 86886-86901, 2020.

[2] Gupta, Rishi, Akhilesh Kumar Sharma, Oorja Garg, Krishna Modi, Shahreen Kasim, Zirawani Baharum, Hairulnizam Mahdi, and Salama A. Mostafa. "WB-CPI: Weather based crop prediction in India using big data analytics." vol. 9, pp. 137869-137885, 2021.

[3] Rashid, Mamunur, Bifta Sama Bari, Yusri Yusup, Mohamad Anuar Kamaruddin, and Nuzhat Khan. "A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction." vol 9, pp.63406-63439, 2021.

[4] Alebele, Yeshanbele, Wenhui Wang, Weiguo Yu, Xue Zhang, Xia Yao, Yongchao Tian, Yan Zhu, Weixing Cao, and Tao Cheng. "Estimation of Crop Yield from Combined Optical and SAR Imagery Using Gaussian Kernel Regression." vol 14, pp.10520-10534, 2021.

[5] Raja, S. P., Barbara Sawicka, Zoran Stamenkovic, and G. Mariammal. "Crop Prediction Based on Characteristics of the Agricultural Environment Using Various Feature Selection Techniques and Classifiers." vol 10, pp. 23625-23641, 2022.

[6] Bondre, Devdatta A., and Santosh Mahagaonkar. "Prediction of crop yield and fertilizer recommendation using machine learning algorithms." vol 4, no. 5, pp.371-376, 2022.

[7] Mariammal, G., A. Suruliandi, S. P. Raja, and E. Poongothai. "Prediction of land suitability for crop cultivation based on soil and environmental characteristics using modified recursive feature elimination techniques with various classifiers." vol 8, no. 5, pp.1132-1142, 2021.

[8] Karthikeya, H. K., K. Sudarshan, and Disha S. Shetty. "Prediction of crops using KNN algorithm."vol 5, pp.1422-1424, 2021.

[9] Narasimhamurthy, V., and P. Kumar. "Rice crop yield forecasting using random forest algorithm."vol 5, pp.1220-1227, 2021.

[10] Domingues, Tiago, Tomás Brandão, and João C. Ferreira. "Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey." vol 12, no. 9, pp.1350, 2021.

[11] Nandhini, S., and K. Ashokkumar. "Machine learning technique for crop disease prediction through crop leaf image."vol 16, no. 2, pp.149-158, 2022

[12] Anitha, A. "Disease prediction and knowledge extraction in banana crop cultivation using decision tree classifiers." vol 20, no. 1, pp. 107-120, 2022.

[13] Tatli, Sana, Esmaeil Mirzaee-Ghaleh, Hekmat Rabbani, Hamed Karami, and Alphus Dan Wilson. "Prediction of Residual NPK Levels in Crop Fruits by Electronic-Nose VOC Analysis following Application of Multiple Fertilizer Rates." vol 12, no. 21, pp. 11263, 2022.

[14] Hossain, M. A., and M. N. A. Siddique. "Online Fertilizer Recommendation System (OFRS): A Step Towards Precision Agriculture And Optimized Fertilizer Usage By Smallholder Farmers In Bangladesh: Online fertilizer recommendation." vol 1, no. 4, 2020.

[15] Alghamdi, Rashad S., and Larry Cihacek. "Do post-harvest crop residues in no-till systems provide for nitrogen needs of following crops?." vol 114, no. 1, pp.835-852, 2022.