Journal of Harbin Engineering University
ISSN: 1006-7043
Vol 45 No. 01
January 2024

Farmer Assistant Fertilizer Prediction System

Anil Kumar B.[1], Anitha J.[2]
[1]Research Scholar, Department of CSE
Visvesvaraya Technological University, Belagavi, Karnataka, India
[2]Professor, Department of CSE
RV Institute of Technology and Management
Visvesvaraya Technological University, Belagavi, Karnataka, India

Abstract: Agriculture is an essential sector in India's economy, with a large contribution to its GDP (Gross Domestic Product) and playing a crucial role in the global economy. The agricultural sector in India is a significant source of employment, with approximately 60-70% of the workforce dependent on it. India has one of the largest arable lands in the world, second only to the United States. The reason behind India's significant dependence on the agricultural sector is the presence of fertile soil and numerous sources of water for irrigation. As the world's population continues to grow, it becomes imperative to ensure food security, making agriculture even more crucial. The production and regulation of agricultural yields determine the extent of food security within the nation. However, given the limited resources available, meeting the demands of a rapidly expanding population presents a significant challenge. The agricultural sector has witnessed a decline in output attributed to factors such as pest infestations, inadequate fertilizer supply, and deteriorating soil fertility. This has further underscored the importance of fertilizers in modern agriculture. Fertilizers primarily consist of nitrogen, phosphorus, and potassium. The aim of this project is to ensure that these components are present in the required amount for better crop yield.

Index Terms—Farmer Assistant, Fertilizer recommendation, Random Forest, Food security.

I. Introduction

India, being a developing nation, agriculture in India is unstable and our farmers are at the forefront of the sector. The Indian economy relies on agriculture for its growth. Technology has to be used extensively in agriculture to overcome the challenges faced by the sector. It can facilitate farmers to analyze weather conditions precisely, reduce waste, enhance the results and increase their profitability. Farmers encounter numerous issues which consist of plants getting affected with diseases, the soil being insufficiently nourished. All these elements lessen the general yield.

The macro-vitamins utilized by plants are Phosphorus, Potassium and Nitrogen. Nitrogen(N) is accountable for the development of the leaves in the plant, Phosphorus(P) for root, fruit, and flower and Potassium(K) allows proper functioning of xylem and is responsible for growth of stem. With the knowledge of nutrients present in fertilizer the system can aid a farmer pick out an appropriate fertilizer. India is a rural nation whose economy is heavily dependent on progress of agricultural yield and integrated agro-industrial elements. It is currently moving rapidly in terms of technological development. Smart farming is revolutionizing agriculture in India.

The yield of crops is determined by the characteristics of the soil and the surrounding environmental conditions. The reason behind this is the inadequate utilization of technology and casual attitude about evaluating the requirements for achieving a better yield which causes problems in agriculture including crops affected by pests, insects, diseases and decreased revenue. In order to address this issue, it is necessary to select the appropriate fertilizer for a specific crop. The use of fertilizers enables farmers to enhance the rate of crop productivity. In the proposed method, different feature engineering methods in conjunction with classification algorithms will be employed to estimate the appropriate fertilizer for a given crop. Since using technology to meet the requirements is the need of the hour, several machine learning techniques can be utilized to perform this process. The proposed system recommends fertilizers to enhance the crop yield analyzed by the soil conditions like Moisture, Temperature, Humidity, Soil Type, Crop Type, Potassium, Phosphorus and Nitrogen.

The main objective of the proposed system is to recommend fertilizer for improving the yield of crops for the Karnataka state. For this, the nutrients of the soil have to be analyzed and the

data gathered is used for recommending fertilizer. The primary nutrient content in fertilizer is Nitrogen, Potassium and Phosphorus, along with this, other parameters such as Moisture, Temperature, Soil Type, Humidity, Crop Type, Potassium, Phosphorus and Nitrogen are considered in this approach. This system will assist the farmers to get good yield and increase the crop produce.

II. Literature Survey

Jeevaganesh et al., [1] predicts fertilizer required and expected crop yield using AdaBoost and Random Forest Algorithm on crops like rice, orange, lentils, mango, maize, black gram, banana, papaya, apple, and grapes. The dataset used involved 6 parameters for fertilizer prediction, moisture, temperature, humidity, NPK values, crop type and soil type. The crop yield was predicted based on state, district, season, crop, area and production. An accuracy of 95% was achieved for fertilizer recommendations and 82% for crop yield prediction.

Chougule et.al., [2] utilized government-provided data on yearly crop production in various regions and maintained this knowledge base using ontology data property values and APIs. Their system recommends suitable crops for a specific area based on the history of the crop yield for the last three years and suggests appropriate fertilizers for particular crops based on soil nutrients. To make these recommendations, the

system employs a random forest algorithm for crop selection and K-means clustering to predict the most appropriate fertilizer for each crop according to the NPK values of the soil. The decision criteria for crop selection take into account both the quantity of production and the market value of the crop in a particular region. The parameters considered for the implementation of the random forest algorithm are state, season and district, while the input to the clustering algorithm is the type of the crop and the composition of the soil-nitrogen (N), potassium (K) and phosphorus (P). The algorithm makes clusters of nearby fertilizers using Euclidean distance, which is the difference between the NPK values. The system recommends a fertilizer that has the least distance in the cluster to the farmer.

V Suma et al. [3], developed a knowledge-based system which included a user interface, working memory, general and domain knowledge, and an inference engine. Convolutional neural network (CNN) was employed to recognize leaf diseases and suggest appropriate pesticides. The use of CNN led to 99.32% accuracy in disease identification.

P.S.Nishant et al., [4], proposed a system which employed three datasets including data for crop recommendation, yield prediction, and crop prices. Various machine learning models that use algorithms like decision tree, random forest, Lasso regression, XGboost, and Ridge regression, were applied on the data for prediction and analysis. In Crop Recommendation, XGBoost is found to be the most accurate with 95% accuracy. Meanwhile, for Crop price and Crop yield prediction, algorithms such as random forest regression are utilized, achieving an accuracy of 92% and 91% respectively.

Wickramasinghe et al., [5], proposed a model whose primary objective was to suggest the most appropriate crop depending on the soil fertility of a particular field and recommend a plan to reduce the quantity of fertilizers used for the specified crops. To accomplish this, the Multi-Class SVM is applied for optimizing the amount of N, P and K application for each crop, taking into account various common fertilizers that are applied in dry and wet regions, availability of nutrients in the soil, and crop nutrient requirements. A mobile application is developed to provide a list of the most suitable crops depending on the conditions of the soil, and the customer can search for a specific crop to determine its suitability for the soil.

Gupta. S et al., [6] developed a fertilizer recommendation system where the user enters N,P,K values and POST requests are made to the flask API. HTTP response is sent and the user gets a recommendation for fertilizer. The response also sends other details such as the best crop farmer can grow in that area. Another system for disease detection is developed where the user has to upload an image and the model recognizes the disease and remedies.

Manasi Jadhav et al.,[7] proposed a system which performs weather forecast, crop recommendation and fertilizer prediction. It was mainly in Maharashtra region and three algorithms were compared - KNN, Naive Bayes and Decision tree, out of which Decision tree gave better results. The accuracy obtained was around 90%.

In [8], three forecasting techniques, namely ARMAX, SARIMA and ARMA were employed and their performance evaluated. The model which proved to be very effective was deployed to measure temperature and rainfall. This was then

used, along with a fuzzy logic model to analyze crop yield. Attributes used were temperature and rainfall.

A model was developed by Schwalbert et al. [9] to predict soybean yields in Brazil utilizing LSTM together with considering the weather data and satellite imagery. The efficacy of a few distinct algorithms, namely random forest, multivariate regression and LSTM, was compared in estimating the yield produced by soybean. Also, land surface temperature, NDVI, precipitation, and EVI were considered. LSTM outperformed the other algorithms.

Gosai, D., [10] proposed a system which suggests growth of suitable crops. Soil testing was done with the help of sensors using IOT and Machine learning. The sensors to measure temperature of the soil, moisture of the soil, pH of the soil, NPK, have been deployed for analyzing soil moisture, soil pH levels, humidity, temperature as well as nutrient contents in soil such as potassium, nitrogen and phosphorus are used and data is stored on a microcontroller. The analysis was performed utilizing Support Vector Machine, Decision Tree, Random Forest, Naive Bayes and XGBoost, and their accuracies were compared. Dataset was obtained from Kaggle and crops such as Rice, Maize, Chickpea among others were used.

Gaikwad et al., [11] developed a system which monitors the soil parameters like soil moisture, temperature and air humidity in real time using arduino based IoT devices. A steel probe is used to collect soil samples which is later used by the sensors to collect data. An application is used to collect the soil parameters measured using sensors and this data is sent to the server to analyze using a web-based technology.

Akhter et al., [12] developed a model for the forecasting of Apple disease among apple orchards based on the Kashmir Valley region using Machine Learning in IoT and data analytics. The system used WSN/IoT nodes to collect data and six sensors were deployed in each node to cover an area of orchard. The collected data is analyzed to predict the presence of apple scab using linear regression.

Bondre et al., [13] proposed and implemented a system that can be used to make predictions about the crop produced by analyzing past information available. Random Forest and Support Vector Machine methods were used to suggest fertilizer for a certain crop. The focus was on creating a model that can be utilized to forecast crop produce.

Sivakumar et al., [14] created a model for optimizing the amount of fertilizers added to the crops by utilizing the Fruit Fly optimization algorithm. This algorithm is a well-known optimization technique inspired from biological processes to address intricate optimization problems. The research was performed in the Coimbatore region of Tamil Nadu state in India. The outcomes of the FFO analysis have demonstrated that the excessive application of fertilizers can be effectively reduced through appropriate optimization planning, thus mitigating the harmful consequences that chemical fertilizers have on agricultural, soil, and the ecosystem.

Klompenburg et al., [15] conducted a Systematic Literature Review (SLR) to gather along with analyzing various features and algorithms employed in previous studies on predicting crop yield. Based on the analysis performed, the features most frequently utilized in these models are temperature, rainfall and soil type while the Artificial Neural Network is the most commonly employed algorithm. The supplementary analysis conducted indicates that CNN is the predominant deep learning algorithm utilized in the studies conducted, with DNN and LSTM also being frequently used. Chiplonker et al.,[16] developed a system that predicts the yield of crop and also provides the alternate crops according to the given condition. Random Forest regression and Linear regression algorithms were used and 94% accuracy was
obtained.

Senthil Kumar Swami Durai et al.,[17] developed a model that recommends crop, identifies weed, and performs pest identification and estimates the cost. The algorithms used for crop prediction gave 90% accuracy. User provides the detail of pH and average rainfall to the website, latitude and longitude of the present location is extracted through web scraping and it is used as input to weather API. The model suggests a suitable crop by analyzing the temperature and humidity data collected from the current location, which is provided by the user.

Hampannavar V. Bhajantri et al.,[18] proposed a work that identifies deficiency of nitrogen in chili plants and recommends quantity of fertilizer consumption by these plants. Histogram analysis and image processing is utilized to classify plants into healthy and nitrogen deficient plants and also the affected region is determined.

III. Proposed System

Through this paper, a system is proposed for recommending fertilizers in a particular area considering the existing contents of soil. Parameters- Moisture, Temperature, Soil Type, Humidity, Crop Type, Potassium, Phosphorus and Nitrogen are considered. The crops to be considered are - Barley, Cotton, Groundnut, Maize, Millets, Oil Seeds, Paddy, Sugarcane, Tobacco, Wheat, and Pulses. The common fertilizers- Urea, DAP, 14-35-14, 28-28, 17-17-17 and 20-20 are included. The fertilizer is recommended based on the soil types, soil nutrients, and various mentioned parameters for the cash crop selected by farmers. Dataset collection is done for nutrients required to grow cash crops, and recommending a fertilizer based on the deficient nutrients. The dataset has been imported from Kaggle.

Data preprocessing includes cleaning and transforming the categorical attributes into numerical features using one-hot encoding. Cleaning data involves removing any errors or inconsistencies in the data. This can involve fixing typos, removing duplicates, and dealing with missing values. Missing values can be problematic because they can cause errors or biases in the model. The missing values present in the dataset were eliminated by removing the rows or columns that contained them. Null values can cause errors in the model, so it is important to handle them appropriately.

The categorical attributes were encoded using one hot encoding. One hot encoding is a method used to transform categorical values to a format that is usable by the algorithms. This requires creation of a binary vector for all the categories in the variable. This makes it easier for the machine learning model to process the data.

Once the data had been encoded, it was partitioned into separate training and testing datasets. The dataset was partitioned into test and train with the split ratio of 80/20.

Standard Scalar is a pre-processing technique that scales the features of a dataset to have zero mean and unit variance. This is done by taking the difference of mean and dividing by the standard deviation for all the features. This can help improve the performance of some machine learning algorithms by ensuring that all features are on the same scale Data augmentation was done to increase the size of the dataset using Variational Autoencoders. Models were developed using SVM, ANN, , Adaboost, Random Forest (with the use of standard scaling).

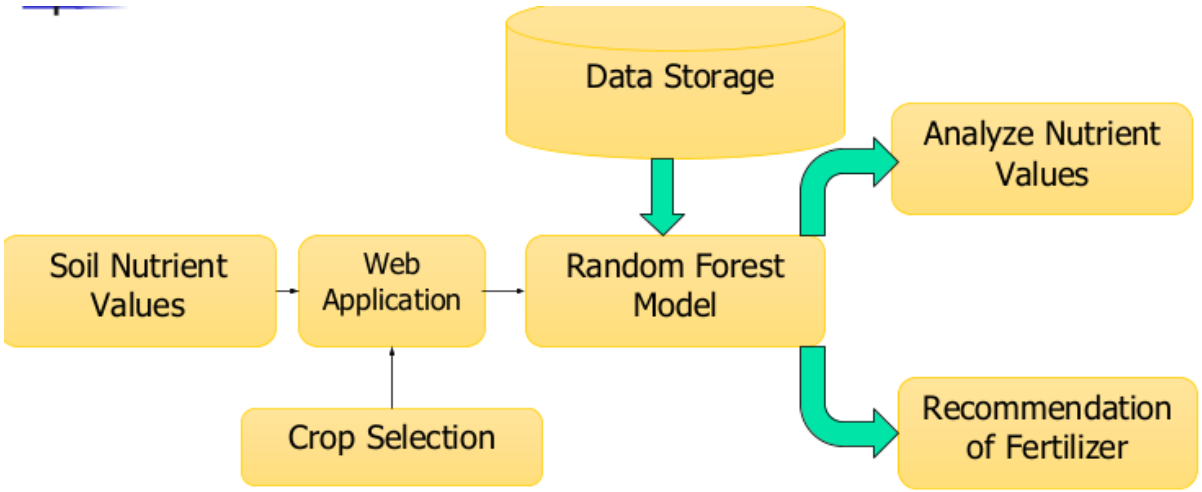

Fig. 1. Block diagram of proposed system

IV. Algorithms

A. Support Vector Machine

Supervised machine learning algorithm, SVM (Support Vector Machine) has the capability to be used for tasks such as classification or regression. In problems such as binary classification, SVM algorithms are designed to locate the optimal hyperplane that effectively divides the two classes within the feature space.

B. Artificial Neural Network

In ANN, information is processed through layers of interconnected nodes or neurons. Each neuron receives input from other neurons and a mathematical function is applied to generate an output. The output of a layer of neuron is input to

the next layer, and this process proceeds until the network produces an output.

C. Adaptive Boosting

AdaBoost is a widely used algorithm in machine learning. It is used to combine multiple weak or base classifiers into a stronger classifier. The basic idea behind AdaBoost is to assign higher weights to the misclassified examples during the training process, so that the subsequent classifiers can learn to focus on the harder-to-classify examples.

D. Random Forest with Standard Scaling

A Standard Scaler and a Random Forest Classifier are combined in this method. The GridSearchCV is then used to find the best hyperparameters for the Random Forest Classifier based on the given parameter grid. The pipeline is then fit on the training data, and used to make predictions on the test data. stronger classifier. The basic idea behind AdaBoost is to assign higher weights to the misclassified examples during the training process, so that the subsequent classifiers can learn to focus on the harder-to-classify examples.

g. Random Forest with Standard Scaling

A Standard Scaler and a Random Forest Classifier are combined in this method. The GridSearchCV is then used to find the best hyperparameters for the Random Forest Classifier based on the given parameter grid. The pipeline is then fit on the training data, and used to make predictions on the test data.

V. Experiments, Results And Discussion

There were six algorithms used and its accuracy compared. Random Forest with standard scaling gave 97.55% which was higher accuracy compared to the other algorithms.

ANN and SVM models did not perform well as they struggled to model the non-linear interactions between the input features. Random forest and Adaptive boost perform with a higher accuracy as they are more capable in handling nonlinear interactions.

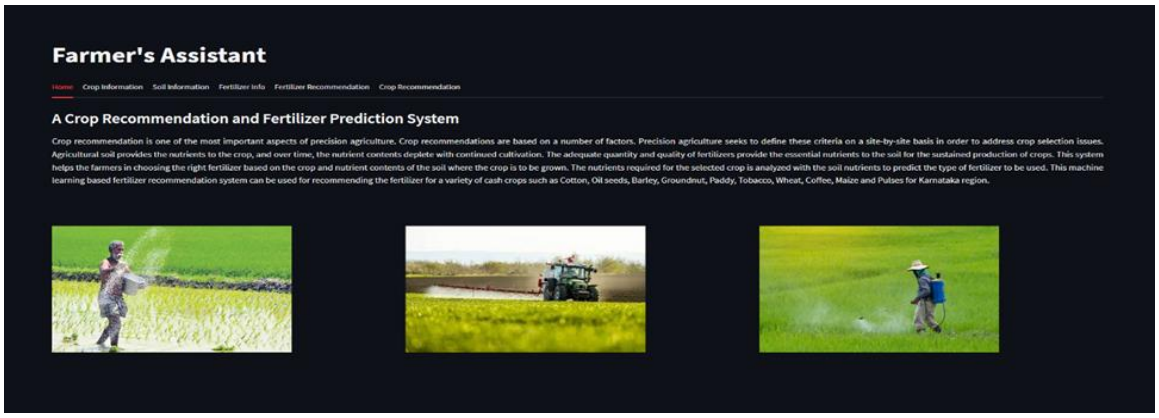

Fig 2: Farmer's assistance

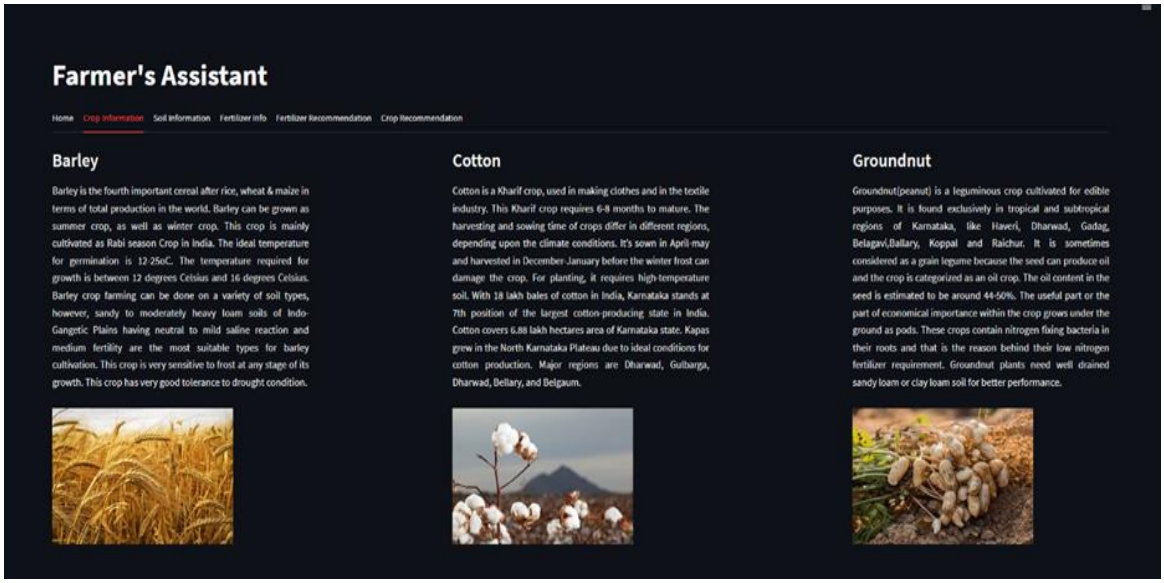

Fig 3: Crop Information

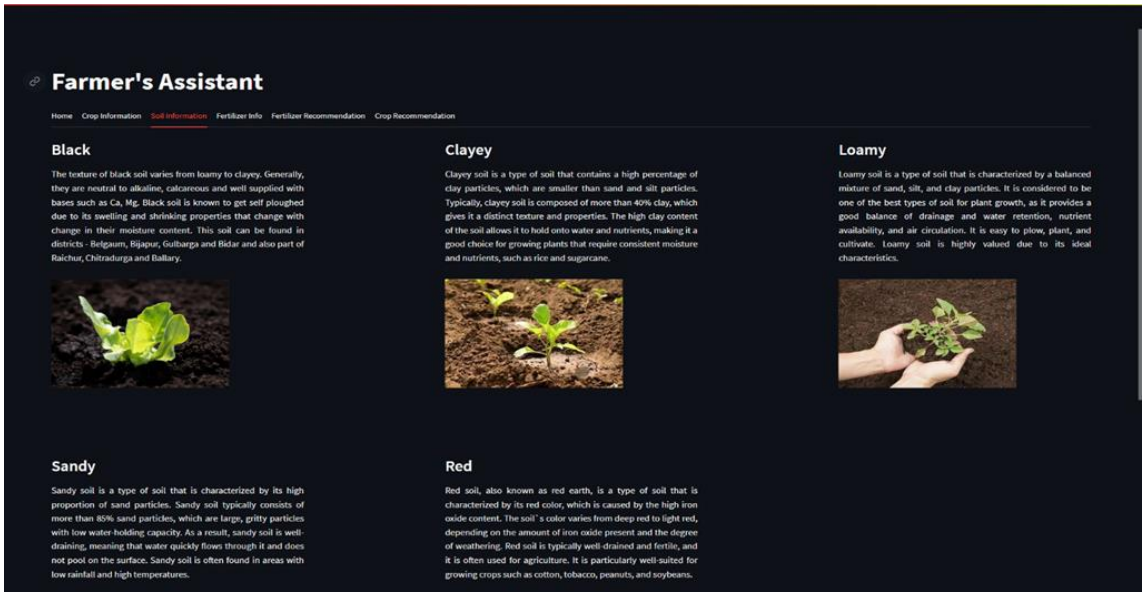

Fig 4: Soil Information

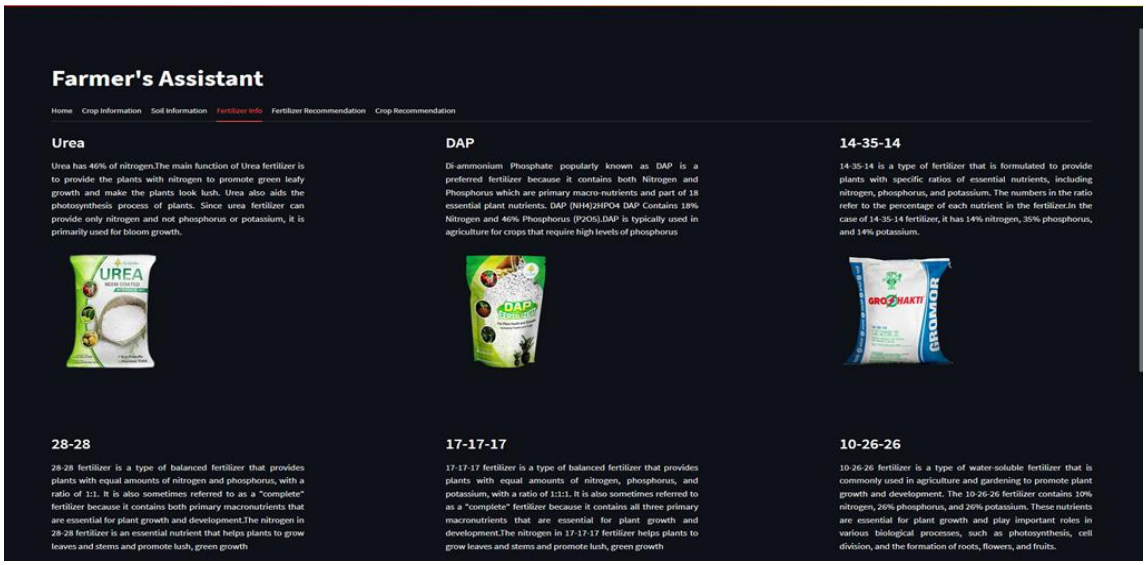

Fig 5: Fertilizer Information

VI. Conclusion

Meeting the needs of a rapidly expanding population with limited resources is a significant challenge. To ensure food security, it is crucial for the nation to maintain an adequate agricultural yield. This can be a daunting task, given the growing demand for food. Since agriculture is a very important sector, especially in India, it is critical to strengthen this sector.

These approaches will enable farmers to understand the type of fertilizer that can be used for a particular soil and to obtain better crop yield. It will also help improve the soil fertility and reduce the damage caused to the crop due to inappropriate usage of fertilizers.

Six common fertilizers are considered namely Urea, DAP, 14-35-14, 28-28, 17-17-17 and 20-20. and the crops are Barley, Cotton, Groundnut, Oil Seeds, Millets, Maize, Paddy, Sugarcane, Tobacco, Wheat, and Pulses.

References

[1] Jeevaganesh, R., Harish, D. and Priya, B., 2022, April. A Machine Learning-based Approach for Crop Yield Prediction and Fertilizer Recommendation. In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1330-1334). IEEE.

[2] Chougule, A., Jha, V.K. and Mukhopadhyay, D., 2019. Crop suitability and fertilizers recommendation using data mining techniques. In Progress in Advanced Computing and Intelligent Engineering (pp. 205-213). Springer, Singapore.

[3] V. Suma, R. A. Shetty, R. F. Tated, S. Rohan and T. S. Pujar, "CNN based Leaf Disease Identification and Remedy Recommendation System," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 395-399, doi: 10.1109/ICECA.2019.8821872.

[4] P. S. Nishant, P. Sai Venkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.

[5] Wickramasinghe, C.P., Lakshitha, P.L.N., Hemapriya, H.P.H.S., Jayakody, A. and Ranasinghe, P.G.N.S., 2019, December. Smart Crop and Fertilizer Prediction System. In 2019 International Conference on Advancements in Computing (ICAC) (pp. 487-492). IEEE.

[6] Gupta, S., Chopade, A., Jain, N. and Bhonde, A., 2022. Farmer's Assistant: A Machine Learning Based Application for Agricultural Solutions. arXiv preprint arXiv:2204.11340.

[7] Jadhav, M., Kolambe, N., Jain, S. and Chaudhari, S., 2021, May. Farming made easy using machine learning. In 2021 2nd International Conference for Emerging Technology (INCET) (pp. 1-5). IEEE.

[8] Bang, S., Bishnoi, R., Chauhan, A.S., Dixit, A.K. and Chawla, I., 2019, August. Fuzzy logic based crop yield prediction using temperature and rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1-6). IEEE.

[9] Schwalbert, R.A., Amado, T., Corassa, G., Pott, L.P., Prasad,
P.V. and Ciampitti, I.A., 2020. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Agricultural and Forest Meteorology, 284, p.107886.

[10] Gosai, D., Raval, C., Nayak, R., Jayswal, H. and Patel, A., 2021. Crop Recommendation System using Machine Learning.

[11] Gaikwad, S.V., Vibhute, A.D., Kale, K.V. and Mehrotra, S.C., 2021. An innovative IoT based system for precision farming. Computers and Electronics in Agriculture, 187, p.106291.

[12] Akhter, R. and Sofi, S.A., 2022. Precision agriculture using IoT data analytics and machine learning. Journal of King Saud University-Computer and Information Sciences, 34(8), pp.5602-5618.

[13] Bondre, D.A. and Mahagaonkar, S., 2019. Prediction of crop yield and fertilizer recommendation using machine learning algorithms. International Journal of Engineering Applied Sciences and Technology, 4(5), pp.371-376.

[14] N. Sivakumar, T. Amudha and N. Thilagavathi, "Development of a Novel Bio Inspired Framework for Fertilizer Optimization," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019, pp. 175-181

[15] Van Klompenburg, T., Kassahun, A. and Catal, C., 2020. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, p.105709.

[16] Chiplonker, Disha. “Crop yield prediction and alternate crop prediction using ML.” (2021).

[17] Senthil Kumar Swami Durai, Mary Divya Shamili, Smart farming using Machine Learning and Deep Learning techniques,Decision Analytics Journal,Volume 3,2022,100041,
ISSN 2772-6622

[18] K. Hampannavar, V. Bhajantri and S. G. Totad, "Prediction of Crop Fertilizer Consumption," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5.