Smart Agriculture Farming Using Harvestify Web App

Sachin Kapoor, Ishika Aggarwal, Anshu Kumar Ray
Department of Computer Science and Engineering, Sharda University, Greater Noida, India
Email: 2020001745.sachin@pg.sharda.ac.in,
2020549602.ishika@pg.sharda.ac.in,
2020423324.anshu@pg.sharda.ac.in

ABSTRACT
In India, the farming sector is highly vital and crucial for economic growth and jobs. In India, agriculture provides a living for approximately 48% of the people. Agriculture provides opportunities for village people to work and contributes to the development of a country like India on a big scale, as well as a boost to the economy.

The research aims to assist farmers in determining the quality of the soil and assessing its many parameters, as well as recommending crops and fertilizers depending on the outcomes acquired through a machine learning technique. To improve the effectiveness of the Crop Recommendations Systems and Fertilizer Recommendation System, the system employs a number of Classification techniques. The assigned soil and fieldwork to anticipate a list of crops that is suited for the soil, as well as knowledge on minerals that are insufficient in the soil again for a crop in question. As a result, the user is free to choose which crop to plant. As a result, the approach aids dilettante farmers in gaining information. This research uses soil and PH data as inputs and uses a website to forecast which crops are suited for the soil and which fertilizers can be used as a remedy.

Keywords **-** Agriculture, Crop Recommendation, Ensembles Model, Fertilizer Recommendation, Machine Learning, Plant Disease Detection, Recommendation System, Smart Farming.

I. INTRODUCTION

In India, the agriculture industry is extremely vital and crucial for economic and social development and jobs. In India, the agricultural sector provides a living for almost 48% of the population. As per the 2019-2020 economic survey, an Indian farmer's median wage in 16 states is Rupees. 2500/-.Most of the Indian population depends on agriculture for their livelihood. Agriculture gives an opportunity of employment to the village people to develop a country like India on large scale and give a push in the economic sector.

The majority of farmers face the problem of planting an inappropriate crop for their land based on a conventional or non-scientific approach. This is a challenging task for a country like India, where agriculture feeds approximately 42% of the population. And the outcomes for the farmer of choosing the wrong crop for land is moving towards metro city for livelihoods, suicide, quitting the agriculture and give land on lease to industrialist or use for the non-agriculture purpose. The outcome of wrong crop selection is less yield and less profit. If the family is fully dependent on farming, so, it is very difficult to survive. Nowadays most farmers don't want their children working in agriculture. Instead, they prefer to settle in the metropolis city and work in different fields. Machine learning, which is one of the applications of Artificial Intelligence, is being used to implement the proposed system. Crop recommendation is going to recommend you the best crop you can grow in your land as per the soil nutrition value and along with as per the climate in that region.

And recommending the best fertilizer for every particular crop is also a challenging task. And the other and most important issue is when a plant gets caught by heterogeneous diseases that effect on less amount of agriculture production and compromises with quality as well. To overcome all these issues this recommendation has been proposed [9]. Nowadays a lot of research and work is being implemented in the smart and modern agriculture domain. Crop recommendation is characterized by a soil database comprised of Nitrogen, Phosphorus, potassium in addition to climate conditions such as

Table 1: Dataset of Crop Recommendation (First 10 rows)

Nitrogen (N)	Phosphorus (P)	Potassium (K)	Temperature	Humidity	pH	Rainfall	Label
90	42	43	20.87974371	82.00274423	6.50985292	202.9355362	Rice
85	58	41	21.77046169	80.31964408	7.038096361	226.6555374	Rice
60	55	44	23.00445915	82.327629	7.840207144	263.9642476	Rice
74	35	40	26.49109635	80.15836264	6.980400905	242.8640342	Rice
78	42	42	20.13017482	81.60487287	7.628472891	262.7173405	Rice
69	37	42	23.05804872	83.37011772	7.073453503	251.0549998	Rice
69	55	38	22.70883798	82.63941394	5.70080568	271.3248604	Rice
94	53	40	20.27774362	8289408619	5.718627178	241.9741949	Rice
89	54	38	24.51588066	83.5352163	6.685346424	230.4462359	Rice

rainfall and Ph level. For fertilizer recommendation Database characteristics are Nitrogen, Phosphorus, Potassium, and crop. To improve crop productivity and best fertilizer use a smart recommendation system has to be built using the ensembles technique [10]. The ensembles technique is used to build a recommendation model that combines the prediction of multiple machine learning. Models to recommend the right crop based on soil value and the best fertilizer to use.

II. Literature Survey

Nowadays more and more research is going in the field of Agriculture domain. Identify what is the challenge in Indian Farming and come up with new solutions to help farmers.

The paper [1], [2], and [5] mainly throw light on the recommendation of the crop which crop to grow according to your soil nutrition value and weather condition. In this paper, the author starts from the very basics of smart farming and slowly moves towards developing a model which will help the farmer to grow crops suitable to the soil along with the weather conditions.

The Machine Learning algorithm used for crop recommendation in this paper is SVM, Decision tree, KNN, ensembles. The conclusion is that the ensembles model is the best suitable for the crop recommendation model [6].

In the "Improved Segmentation Approach for Plant Disease Detection" paper [4] the author uses the approach in which he tried different-different machine learning models based on the micro and macronutrients like Nitrogen, Phosphorus, pH level, Rain value in mm to predict the best suitable fertilizer for the selected crop. The performance matrix of the classification algorithm is compared based on accuracy and execution time.

"Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review" paper [3] presents a vivid representation of plant disease detection using different-different techniques and tries to enhance the yield. The author has tried to solve this issue by using the neural network state of the art technique.

In the Smart Farming Prediction Using Machine Learning" paper [7] the needs and preparation needed to develop a modelling framework for smart agriculture are covered in this study. It delves into the fundamentals of smart agriculture. The authors begin with the fundamentals of smart agriculture before moving on to constructing a model to support it. This describes a model that uses Precision Agriculture (PA) concepts to reduce variability on small, open fields at the single farmer and crop levels. The model's overall goal is to provide immediate advising service with even the tiniest farmers at the level of his or her lowest crop plot, and use the most available technologies available, such as SMS and email. This model was developed for the situation in Kerala, in which the average retention size is significantly

III. METHODOLOGY

Crop & Fertilizer recommendation: The following is a step-by-step strategy for building the recommendation mode (Fig l):

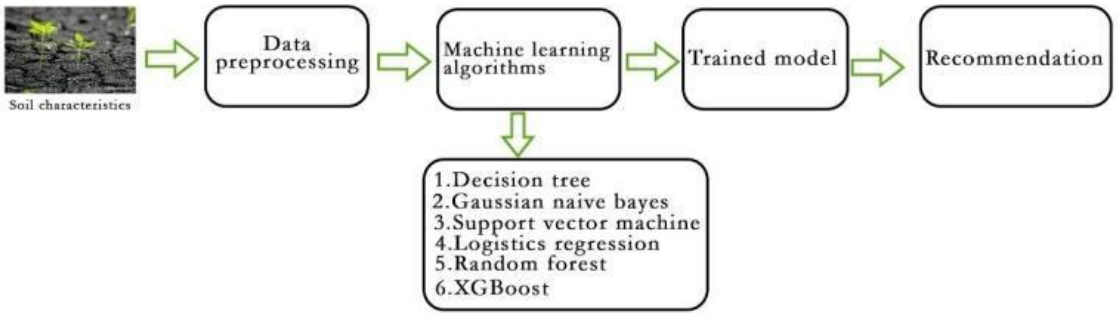

Figure 1: Flow chart of all the steps performed for the classification

quality machine learning model with minimum errors depends upon the quality of data.

Step 2: Preprocessing of the input dataset

In any machine learning project, the most important or the time consuming process is preprocessing the data, In preprocessing step the missing values are being filled using techniques like mean, mode, and median, scaling or transforming values into a certain range, cleaning the data, encoding of categorical data and check for correlation of variable.

Step 3: Exploratory data analysis

In this step, univariate, bivariate, and multivariate analysis are performed to find the hidden patterns in the data and try to understand the data before getting hands dirty into model building. A few examples of univariate analysis plot is PDF and CDF and for the multivariate plot is pair plot, box plot, and Heat map.

Step 1: Load the dataset

The data has been collected this from the "Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique" paper with dataset of crop recommendation (Table 1). A good

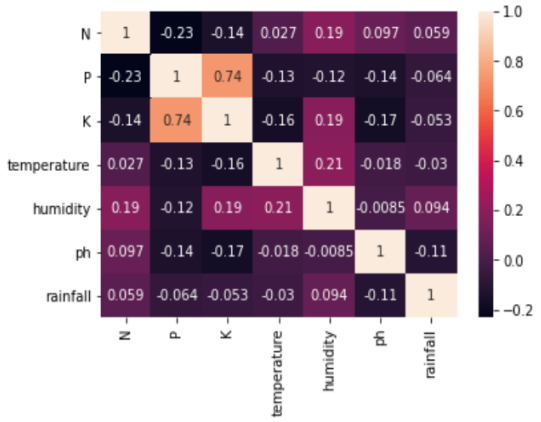

Figure 2: Correlational Metrics on the crop recommendation dataset

Step 4: Splitting into training and testing data

In this step, the preprocessed dataset is split into training and testing based on 80:20 ratios, which means 80% of the data is used for training the dataset and 20% is used for testing on the unseen dataset and cross-validation to find optimal hyper parameter.

Step 5: Building classification model on the training dataset

In this step, the training dataset is given to the individual classifier and train the model on top of it.

The model to build the Crop recommendation system is

Decision Tree: A supervised machine learning algorithm known as a tree structure can be used for both classification and regression [11]. The structure of a decision tree is similar to that of a flowchart, with characteristics and class labels represented by a tree.

Gaussian Naive Bayes: Gaussian Naive Bayes is a machine learning approach that is both basic and straightforward. The Naive Bayes hypothesis is that characteristics must be independent of one another [12]. The Bayes theorem is used internally in Binary Classification, which is a statistical-based method. The model is known as Gaussian Nave Bayes if the features of the dataset follow a Gaussian distribution.

Support Vector Machine: A machine learning technique, SVM stands for SVM algorithm. First plot each data element in N Dimension space, and then choose a hyperplane that best segregates the two classes with the highest margin in the Linear Kernel [13]. Finding the best hyperplane in a Support vector machine is a difficult task.

Logistic Regression: Regression is a machine learning method that is used to classify data [14]. In logistic regression, examine each data point in n dimensions and attempt to locate a hyperplane that separates the two classes of data.

Random Forest: Random Forest is an ensemble-based machine learning system. Ensemble techniques are a form of technology that allows us to combine separate or similar algorithms to create a strong model [15]. Random forest is a compilation of several decision trees with the greatest depth till the nodes can split with the least variability and bias.

XGBoost: XGBoost is a monitored machine learning method. XGBoost is a widely used gradient boosting implementation of a decision tree-based machine learning method [16]. XGBoost uses a more advanced tree-based machine learning technique.

Accuracy Comparison between all the algorithms (Fig. 3):

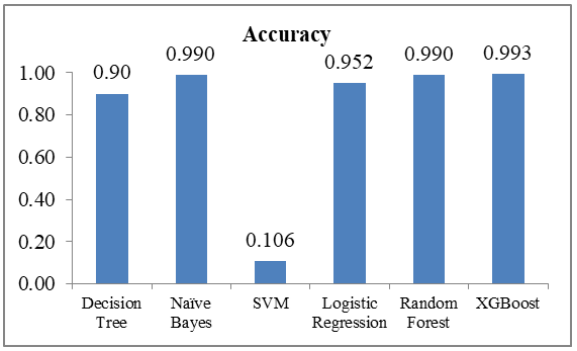

Figure 3: Accuracy comparison chart of different algorithms used in the classification model

Step 6: Testing the data on each of the classifiers

In this step, the query point is given to each and every classifier and obtains the class label from every model

Step 7: Ensemble the individual's classifier output using the majority voting technique

The last step is to get the class label from every classifier after then do majority voting to get the final prediction.

IV. CONCLUSION

Smart Farming system has successfully built that recommends the best crop to grow inland, along with that recommending the best fertilizer to use and plant disease detection which can be easily available and used by farmers in order to take a decision which crop to grow and maximize the yield according to the soil nutrition value and climate in that region.

The model proposed in this paper can be further extended in the future to come in the mobile app to provide recommendations of crop and fertilizer plant disease detection.

The accuracy obtained by using the ensemble technique is 99.01% which is phenomenal well.

REFERENCES

[1] Priyadarshini A, Swapneel Chakraborty, Aayush Kumar, and Omen Rajendra Pooniwala. *"Intelligent Crop Recommendation System using Machine Learning".* https://ieeexplore.ieee.org/document/9418375

[2] Nidhi H Kulkarni, G N Srinivasan, B M Sagar, and N K Cauvery. *"Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique".* https://ieeexplore.ieee.org/document/8768790

[3] Santhosh S. Kumar and B.K. Raghavendra *"Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review".* https://ieeexplore.ieee.org/document/8728325

[4] Md. Arifur Rahman, Md. Mukitul Islam, G M Shahir Mahdee, and Md. Wasi Ul Kabir *"Improved Segmentation Approach for Plant Disease Detection".* https://ieeexplore.ieee.org/document/8934895/authors#authors

[5] Devdatta A. Bondre and Mr. Santosh Mahagaonkar *"Prediction Of Crop Yield And Fertilizer Recommendation Using Machine Learning Algorithms".* https://www.ijeast.com/papers/371-376,Tesma405,IJEAST.pdf

[6] Anantha Reddy Dasari *"Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique".* https://www.irjet.net/archives/V4/i12/IRJET-V4I12179.pdf

[7] S.R.Rajeswari , Parth Khunteta, Subham Kumar,Amrit Raj Singh,Vaibhav Pandey *"Smart Farming Prediction Using Machine Learning"*

[8] Darshan Gandhi *"FarmEasy: Crop Recommendation for Farmers made easy"*

[9] N Nithin Srivatsav *"Recommendation System for Farming :: Iot and ML"*

[10] N. V. S. Sai Teja, P. Sri Gowtham, T. Lakshmi Surekha, P. Gopala Krishna *"Soil Analysis for Fertilizer Prediction using IoT and Machine Learning"*

[11] H. Zhang and R. Zhou, *"The analysis and optimization of decision tree based on ID3 algorithm,"* 2017. https://ieeexplore.ieee.org/document/8321588

[12] A. H. Jahromi and M. Taheri, *"A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features,"* 2017. https://ieeexplore.ieee.org/document/8324083

[13] S. Ghosh, A. Dasgupta and A. Swetapadma, *"A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification,"* 2019. https://ieeexplore.ieee.org/document/8908018

[14] X. Zou, Y. Hu, Z. Tian and K. Shen, *"Logistic Regression Model Optimization and Case Analysis,"* 2019. https://ieeexplore.ieee.org/document/8962457

[15] J. k. Jaiswal and R. Samikannu, *"Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression,"* 2017. https://ieeexplore.ieee.org/document/8074494

[16] L. Sun, *"Application and Improvement of Xgboost Algorithm Based on Multiple Parameter Optimization Strategy,"* 2020. https://ieeexplore.ieee.org/document/9421520

[17] A. Sharma, M. Bhargava and A. Vijay Khanna, *"AI-Farm: A crop recommendation system,"* 2021. https://ieeexplore.ieee.org/document/9421520

[18] K. Taunk, S. De, S. Verma and A. Swetapadma, *"A Brief Review of Nearest Neighbor Algorithm for Learning and Classification,"* 2019. https://ieeexplore.ieee.org/document/90657

[19] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, *"Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction",* International Journal of Computer Science Issues, Volume 9, Issue 3

[20] P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi, and S. G. Koolagudi, *"Tomato leaf disease detection using convolutional neural networks".* In Proc. of the Eleventh International Conference on Contemporary Computing (IC3), pages 15, Aug 2018.

[21] C. G. Dhaware and K. H. Wanjale, *"A modern approach for plant leaf disease classification which depends on leaf image processing".* In Proc. of the International Conference on Computer Communication and Informatics (ICCCI), pages 14, Jan 2017.

[22] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, *"Image texture feature extraction using glcm approach".* In International Journal of Scientific and Research Publications, vol.3