Precision Agriculture Website using Machine Learning

Sumeet Ingole
Department of Computer Science & Engg.
SKN Sinhgad College of Engineering
Pandharpur
sumeet.ingole@sknscoe.ac.in

Mansi Navale
Department of Computer Science & Engg.
SKN Sinhgad College of Engineering
Pandharpur
mansinavale9205@gmail.com

Munira Khan
Department of Computer Science & Engg.
SKN Sinhgad College of Engineering
Pandharpur
munirakhan20112001@gmail.com

Priya Chandole
Department of Computer Science & Engg.
SKN Sinhgad College of Engineering
Pandharpur
priyachandole4@gmail.com

***Abstract*—** Economy of India highly depends on agriculture. Still traditional ways of recommendations are used for agriculture. Currently, agriculture is done based on various approximations of fertilizers quantity and the type of crop to be grown or planted. Agriculture highly depends on the nature of soil and climate. Therefore, it becomes important to make advancement in this field. The project proposes development of an ontology-based recommendation system for crop suitability and fertilizers recommendation. It bridges the gap between farmers and technology. The system predicts suitable crop for the field under consideration based on region of India and type of soil. It provides proper recommendation of fertilizers to the farmers. Fertilizer Recommendation is done based on Nitrogen, Phosphorus, Potassium (NPK) contents of soil and using past years' research data that is stored in Ontology. Along with fertilizer recommendation system also provides suggestions about crop suitability in particular region. Recommendation system uses random forest algorithm and K-means clustering algorithm. The Project also deals with plant leaf disease detection. In agriculture, research of automatic plant disease is essential one in monitoring large fields of plants, and thus automatically detects symptoms of disease as soon as they appear on plant leaves. In this project you will get to know about different types of fertilizers to be used according to the crops and the type of soil. It will also provide you with the information of various diseases plants get affected to and also provide proper pesticides for your crops. So, overall it is going to help you with all of your problems with agriculture.

Keywords— *Machine Learning, SVM, CNN, Random Forest Algorithm,Decision Tree*

I. INTRODUCTION

Agriculture is the main source of income and survival in India for majority population. Agriculture is done from ages. Hence a rich collection of agricultural past data is available. Information Technology can be used to process such a large amount of data and then for recommendation. Various data mining techniques can be used for finding recommendations about crops and fertilizers. Outputs of these techniques can be communicated on the smart phones. This project focuses on the implementation of data mining algorithms which can help in building an efficient recommendation system using available observations data. The paper describes a system which recommends the crops suitable for particular region based on crop yield history of last three years in that region and the fertilizers suitable for specific crop based on soil measurements to farmers. It can help farmers for increasing their crop production. The project shows how information available with government about yearly production in various areas can be used for crop recommendations to farmers. As information represented in the form of ontology can be easily shared and reused, the knowledge base of recommendation system is maintained in the form of ontology. The system uses random forest algorithm for crop recommendation as it works efficiently on huge dataset and can handle missing values. Project describes how k-means clustering can be used for predicting best suitable fertilizer for the crop based on given available NPK content in the soil. Several plant diseases have the potential to bring disastrous economic, social, and ecological consequences. To meet projected demand, worldwide crop production must increase by at least 50 per cent by 2050. Currently, the majority of production takes place in Africa and Asia, where 83 percent of growers are family-run businesses with little to no horticultural experience. As a result, yield losses of more than 50 per cent are common as a result of pests and diseases. The old approach of human analysis by visual inspection is no longer viable for categorizing agricultural diseases. The one of the goal of this project is to identify the most common diseases seen in Plant Leaf. Different types of plants are employed in this research. This system can detect different forms of illnesses in these plants. The user can upload a leaf image, and if the leaf has a disease, the name of the disease will be presented on the user's screen and also the cure for it after pressing the predict button. If there is no disease in the plant leaf after uploading the image, a message reading "There is no disease on the Plant" will be displayed.

II. LITERATURE REVIEW

1. Limin Chuana and Ping Hea proposed a fertilizer recommendation system for wheat in China [1]. Two parameters namely yield response and agronomic efficiency are used by the recommendation system. Limin Chuana and Ping Hea also consider the Nitrogen, Phosphorous and Potassium (NPK) contents for fertilizer recommendation of wheat. It helps to prevent the inappropriate application of fertilizers in wheat production systems in China. Yield response and agronomic efficiency were incorporated as part of the Nutrient Expert for Wheat fertilizer recommendation decision support system.
2. Department of agriculture, government of West Bengal has developed soil test based fertilizer recommendation system (STFRS) for farmers in West Bengal [10]. Information about soil testing laboratories, availability of nutrients and recommendations by experts are provided to farmers through sms service on mobile phones. Smart soil health card provides access to cloud based data on mobile phones. Display of digital soil maps is also an important feature of the system.
3. QUEFTS model was used for calculating soil fertility required for fertilizer recommendation [9]. Web based decision support system for fertilization application on wheat, maize and peanut is provided by Hao Zhang [2] et. al.. It is specifically developed for villages in China. Maps of villages are taken and location specific recommendations are provided using ArcView in ArcGIS. Three parameters are considered for recommendations namely soil measurements, farm production level and target yield for the crop. Three types of databases as system database, spatial database and attribute database are maintained here. Along with other attributes, meteorological information is also considered for decision making.
4. Recommendations for purchasing fertilizers from online portal based on history of past purchases is proposed by Mansi Shinde et. al. [3]. Apriori algorithm is used for this purpose. Fertilizers analysis services is provided by Spectrum Analytics Inc. Washington [4]. It provides recommendations for 250 type of crops. It provides fertilizers schedule along with its quantity at each growth stage of crop under consideration.
5. Precision Fertilization Management Information System (PFMIS) [5] is fertilization recommendation system based on GIS and GPS. ArcGIS is used for maps on soil resources. Recommendations are done by applying data mining techniques on information collected by GIS and GPS. Mansi Shinde el. al. [7] have proposed a crop recommendation and fertilizer purchase system which uses apriori algorithm for recommendation. Based on previous history of fertilizers purchase recommendations are provided. For crop recommendations, they have used market trends data and applied random forest

III. OBJECTIVE

To implement precision agriculture (A modern farming technique that uses research data of soil characteristics, soil types, crop yield data collection and suggests the farmers the right crop based on their site specific parameters to reduce the wrong choice on a crop and increase in productivity).

To solve the problem by proposing a recommendation system through an ensemble model with majority voting technique crop for the site specific parameters with high accuracy and efficiency. To recommend fertilizer on the basis of N, P, K values and crop. To recognize the pest and recommend suitable pesticide for the same.

IV. METHODOLOGY

Data Cleaning and Pre-processing- One of the first steps is to make sure that the dataset we are using is accurate. The dataset should not have any missing values and if the dataset does have missing values, they should be replaced by the appropriate values. The data should also be checked to see if there is a normal distribution for its features. The outliers should be removed. The skew value of the features should be checked and if the features have skewness, then those features should be normalized by using transformations. The dataset which we used had features having skewness in them. To normalize them, we have used quantile transformation on the features of our dataset.

Data Analysis and Visualization- After performing the cleaning and preprocessing of the data, we perform data analysis and visualizations on our dataset. We try to analyze our data more clearly to find any trends or patterns in the dataset. We have created several visualizations of our dataset in order to understand the data properly. We have created bar charts, scatter plots, box plots etc. in order to visualize the data and find if there are any trends or patterns which we can find that will be useful while implementing our project. Feature Selection It is important that we select only those features that will be necessary to determine the type of crop to grow. For this, we have created a correlation matrix that shows the linear relationship of a feature with every other features. If features are highly correlated, then that feature should be dropped, but as we can see in the below matrix that the features are not highly correlated with each other, hence it makes sense not to drop any of them and hence we will be using all of them to predict the type of crop to grow.

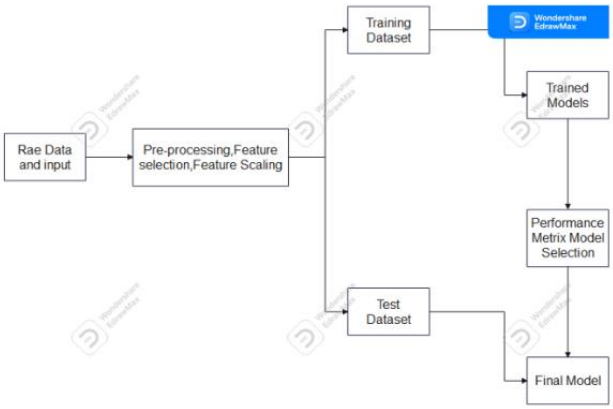

Fig. 1. Flow Chart

Model Building The next step is building the machine learning model. While building the machine learning model, first we need to split our dataset into 2 parts i.e.: training data and test data. We have split the data in the ratio of 70-

30. Taking the training data, we apply our machine learning algorithms on the features of the dataset. We have used 4 machine learning algorithms on our training dataset and the algorithms that gives us the highest accuracy will be selected on the test dataset.

Building a UI In the next step, we have built a UI for a user to input his data so that once he enters the information such as N, P, K values of soil, temperature, humidity, rainfall etc., and the model will process the data and will recommend the appropriate type of crop to be grown in such a condition. Once the user enters the following values and submits the machine learning model will predict the crop that the person must grow.

Machine Learning Algorithms Used Random Forest: Random Forest is a supervised ensemble machine learning algorithm used in both classification as well as regression problems. It contains various decision trees and an average of it is taken so as to give the output. It is based on the concept of bagging wherein multiple decision trees are created and an average of the mistaken so as to give the output. As decision tree are prone to over fitting, random forest is useful in reducing the effect of over fitting and hence giving a more accurate output.

Decision Tree is one of the most popular machine learning algorithms used mostly in classification problems but can also be used for regression type of problems. The working of it is based on a simple technique, wherein a yes/no question is asked and according to the answer the tree is split in smaller nodes. The split of the nodes can either happen by calculating Gini impurity (calculates the measure of impurity) or information gain (calculates the change in the entropy). Decision Trees are prone to over fit and hence this may lead to getting a lower accuracy. This problem can be solved by using random forest algorithm.

Logistic Regression is one of the simplest algorithms in machine learning. It is used for solving classification problems. It uses a sigmoid function to mathematically calculate the probability of an observation and accordingly, the observation is then put into its respective class. While calculating, if the probability of an observation is 0 or 1, a threshold value is decided upon and classes having probabilities above the threshold value are given the value 1 and classes having values below the threshold are given the value 0.

K-Means Clustering is an Unsupervised Learning algorithm, which groups the unlabeled dataset into different clusters. It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabeled dataset on its own without the need for any training. It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters. The algorithm takes the unlabeled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm

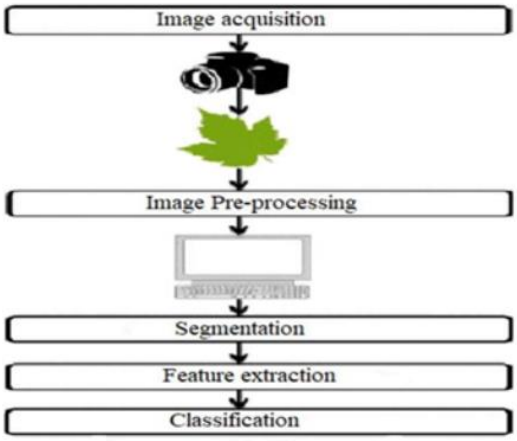

Fig. 2. Block diagram for dieses prediction

V. PROPOSED WORK

A crop recommendation system is essential for optimizing agricultural practices, increasing yield, promoting sustainability, and supporting farmers in making informed decisions. By leveraging data-driven insights and advanced technology, such systems contribute to improved productivity, profitability, and environmental stewardship in the field of agriculture. A fertilizer recommendation system is essential for efficient nutrient management, cost savings, environmental sustainability, soil health preservation, crop-specific customization, and decision support for farmers. A leaf disease detection system is essential for early disease detection, precision disease management, yield protection, resource optimization, disease monitoring and prevention, and decision support for farmers.

VI. RESULT AND ANALYSIS

```
sns.heatmap(df.corr(),annot=True)
```

```
<AxesSubplot:>
```

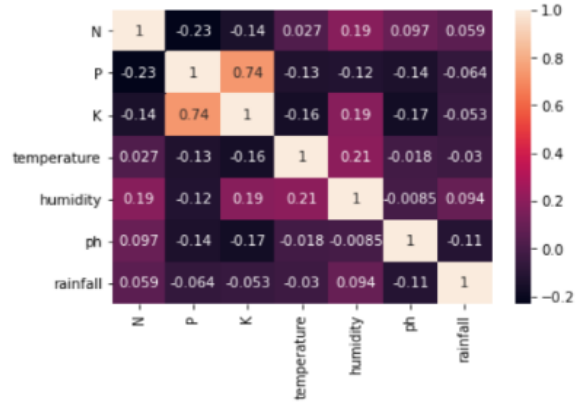

Fig. 3. Heat Map

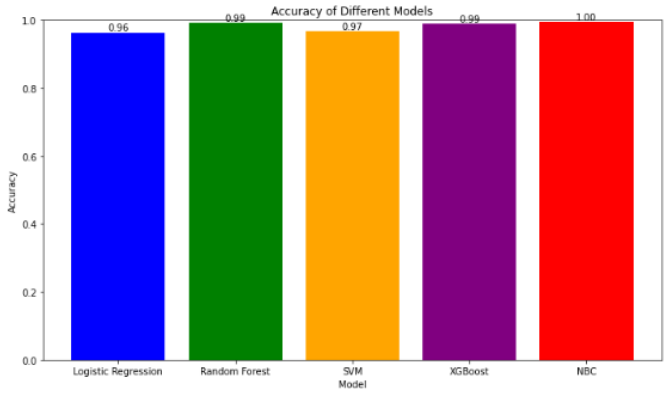

Fig. 4. Accuracy of different algorithms

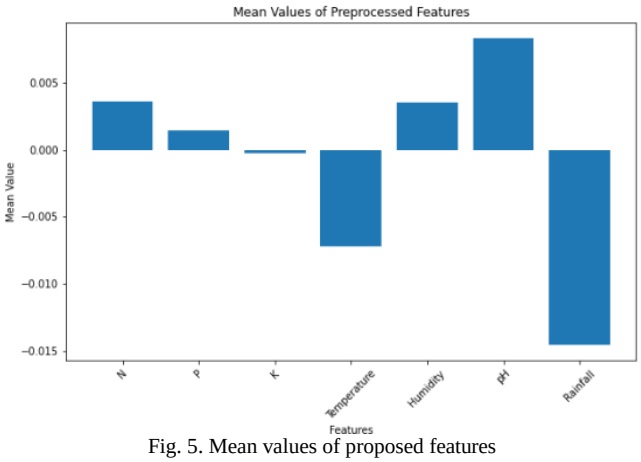

Fig. 5. Mean values of proposed features

CONCLUSION

As our farmers presently are not using technology in their farming practices, they are lagging behind in their farming approaches. Hence, we have made this project so as to encourage farmers to use the current technology instead of relying on old methods. Growing a crop requires a lot of knowledge and insight into many things such as the contents of soil, temperature of the place, pH of soil, proper fertilizers etc. and hence it would make sense for the farmer to adopt to newer technologies as it would make his life easier. Using machine learning algorithms, we will be predicting the right crop to grow, which fertilizer to use and predict what kind of plant leaf disease is present and recommend pesticides for the disease predicted and if the farmer adopts to this technology which would not only make his life easier but also help him in making decisions that do not exploit the environment.

REFERENCES

[1] S. C. S. S. A. Keerthan Kumar TG, “Random Forest Algorithm for Soil Fertility Prediction and Grading using Machine Learning International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November,, 2019.

[2] V. K. J. A Chougule and D. Mukhopadhyay, ” Crop Suitability and Fertilizers Recommendation using Data Mining Techni" Springer Nature Singapore Pte Ltd,, 2019.

[3] S. M. I. K. C. M. Sk Al Zaminur Rahman, ” Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series" 2018 21st International Conference of Computer and Information Technology (ICCIT), 21-23 December,, 2018.

[4] S. N. R. A. P. N. S. Zeel Doshi, ” AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algo- rithms" 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA, 2018).

[5] Z. B. H. et a, ” Feasibility Study on Plant Chili Disease Detection Using Image Processing Technique" ("2012 Third International Conference on Intelligent Systems Modelling and Simulation, pp. 291-296,, 2012).R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[6] G. P. S. . G. R. K. H. M. R. M. Prakash and T. Kaviya, ” Detection of leaf diseases and classification using digital image processing" ("2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS),, 2017).

[7] S. D. Khirade and A. B. Patil, ” Plant Disease Detection Using Image Processing" ("2015 International Conference on Computing Communi-cation Control and Automation, 2015).