International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-12, MAR 2022

# AgroGrow-An Intelligent Crop & Fertilizer Recommendation System using ML

[1]Manasi Jadhav, [2]Ruchika Kale, [3]Amol Kalugade, [4]Rasika Mhabdi
[1,2,3,4]Department of Computer Engineering A.P Shah Institute of Technology Thane, India
[1]mevaja17@gmail.com, [2]kale.ruchika45@gmail.com, [3]arkalugade@apsit.edu.in, [4]rasika@gmail.com

**Abstract— Agriculture plays a significant role in the Indian economy. The majority of Indians rely on agriculture for their living, either directly or indirectly. Agriculture is an important part of India's socio-economic foundation. Failure of farmers to decide on the best-suited crop for the land using traditional and non-scientific methods is a serious issue for a country where approximately 56.6 percent of the population is involved in farming. As a result, it is undeniable that agriculture plays a crucial role in the country. A vast majority of Indian farmers rely on their intuition to determine which crop to plant in a given season. They find solace in simply following ancestral farming standards and practices, unaware of the fact that crop yield is highly dependent on current weather and soil conditions. This leads to change in profession by farmers, quitting the agriculture field, and moving towards urban areas for livelihood. A single misguided or imprudent decision by the farmer can have undesirable consequences on both himself as well as the agricultural economy of the region. To overcome this issue we present an intelligent system, called AgroGrow, which intends to assist the Indian farmers in making an informed decision about which crop to grow considering all the factors like soil characteristics, geographical location, PH level as well as environmental factors such as temperature and rainfall along with how effectively precise quantity of organic fertilizers should be used.**

***Keywords— Agriculture, Crop Recommendation, Machine Learning, organic fertilizers***

## I. INTRODUCTION

Agriculture is one of the major sources of livelihood for about 56% of our nation's population.[1] India is one of the established nations that has agriculture as its primary source of income. Agriculture is one such domain that contributes only around 14% to the GDP but has a considerable amount of impact on the Indian economy. Conventional agricultural practices and techniques pose a lot of issues in terms of efficiency, cost-effectiveness, and resource utilization. There is a requirement for better techniques that can improve the standard of living of the farmers too. Over the years because of globalization, agriculture has evolved by adapting to the latest technologies and techniques for a better standard of living. Among the technologies, precision agriculture is one emerging technology in the field of agriculture.[2] As per the 2016-17, Economic survey the average monthly income of a farmer in 17 states is Rs.1700/- which results in farmer suicides, and diversion of agricultural land for non-agricultural purposes.[7] Besides, 48% of farmers don't want the next generation to take care of their agriculture instead want to settle down in urban areas.[7] The reason behind this is that the farmers often take wrong decisions in the usage of a number of chemical fertilizers that lead to barren lands [3]. The farmer might have purchased the land from others so without previous experience the decision might have been taken. Wrong crop and fertilizer selection will always result in less yield. If the family is fully dependent on this income then it's very difficult to survive.

Both availability and accessibility of correct and up-to-date information hinder potential researchers from working on developing country case studies. With resources in our reach, a system has been proposed to address this problem by providing a predictive understanding of crop sustainability and recommendations based on machine learning models trained considering essential environmental and economic parameters.[4]

A farmer's decision about which crop to grow with a defined quantity of chemical fertilizers is usually clouded by his intuition and other irrelevant factors like making instant profits, lack of awareness about market demand, overestimating a soil's potential to support a particular crop, etc. A misguided decision on the part of the farmer could place a serious burden on his family's financial condition. Maybe this could be one of the many reasons contributing to the endless suicide cases of farmers that are being broadcasted in media daily. In a country like India, where the agricultural sector contributes to approximately 20.4 percent of its Gross Value Added (GVA) [5], such an erroneous judgment would have negative impacts on not just

 DOI : 10.35291/2454-9150.2022.0092 

the farmer's family, but the entire economy of a region.[6] For this reason, we have identified a farmer's dilemma about which crop to grow during a particular season and using a precise quantity of organic fertilizers, as a very grave one.

The need of the hour is to design a system that will provide predictive insights to the Indian farmers, thereby helping them to make an informed decision about which crop to grow and what quantity of organic fertilizer should be used. With this in mind, we propose Agro-Grow - an intelligent system that would consider environmental parameters (temperature, rainfall, humidity, geographical location) and soil characteristics (pH value, soil type, etc) before recommending the most suitable crop and organic fertilizer to the user. In this paper, real-time weather conditions are taken into consideration to design an efficient crop recommendation mechanism.

The proposed system is implemented using machine learning & Artificial Intelligence that allows the systems to learn and evolve automatically without being explicitly programmed. Consequently, the accuracy of the program will be improved without human intervention. Many researchers are researching this field to assist the farmers in the selection discussed as follows, to choose a suitable crop with its various factors like physical, environmental, and economic factors are taken into consideration.

## II. OBJECTIVES

The first objective of this project is to develop a model that will help the farmers to decide which crop will be best suitable for their soil and based on other conditions, that will result in maximum yield for their harvest. The other objective is to propose a system that will give them organic fertilizer recommendations based on the soil nutrient(Nitrogen, phosphorous, potassium) conditions like what the soil lacks or has an excess of and accordingly recommend organic ways to enhance the quality of the soil based on the crops they want to grow.

## III. LITERATURE SURVEY

In this paper by Z. Doshi, five different algorithms are used to suggest the crops for the underlying soil series. They are as follows, Support Vector Machine, Bagged Tree, Adaboost, Naive Bayes, and Artificial Neural Network. Further, the ensemble method is included to provide more accurate results.[9]

In this paper by Priyadharshini A, the proposed system employs two algorithms, linear regression, and neural networks, which are implemented using Pandas, Numpy, Tensorflow, Keras, Scikit-learn, Seaborn, and Python as the programming language. Whereas both the algorithms are based on supervised learning. The results are analyzed and are compared with crop recommendation using K Nearest Neighbour, with cross-validation, Decision Tree, Naive Bayes, and Support Vector Machine concerning accuracy.[7]

In a paper by T.R. Lekhaa, the use of several algorithms like Artificial Neural Network, K-Nearest Neighbors, and Regularized Greedy Forest is demonstrated in selecting a crop based on the prediction yield rate, which, in turn, is influenced by multiple parameters. Pesticide prediction and online trading based on agricultural commodities are also included in the system.[10]

A crop recommendation system for the four crops Rice, Cotton, Sugarcane, and Wheat has been developed in a paper by Nidhi H Kulkarni. The soil dataset is preprocessed initially, and the ensembling approach is then applied in the categorization of the four crops. Random Forest, Naive Bayes, and Linear SVM are the individual base learners employed in the ensemble model. To give the best accuracy, the Majority Voting Technique was employed as part of a combination procedure. The ensembling technique has a precision of 99.91 percent. As a result, the suggested work assists the farmer in making an accurate crop selection for cultivation. This results in an exponential increase in crop productivity, which increases the country's economy.[11]

This paper by Lokesh. K the author focuses on the development of a crop prediction system based on sensor networks that were created with the help of the Internet of Things. The results of the submitted soil samples take a long time to come back from soil testing labs. As a result, the technology claims to assist farmers in obtaining a better crop prediction with no delay in the waiting period [2]. The authors of the research primarily focused on assessing the nitrogen (Nitrogen), phosphorus (Phosphorus), and potassium (K) levels in the soil sample taken for the survey. Based on the data acquired by the sensor network, the suggested method in the research efficiently assesses soil nutrients. This allows for the prediction of the best crop for the soil under test. Farmers must register their NPK sensor with the primary server. From the soil sample and raspberry pi, the NPK extracts the supplement level.[12]

Another intelligent model is described in this paper by Jay Gholap, which enables the prediction of soil properties such as phosphorus content. To obtain high prediction accuracy, the authors apply a variety of classification approaches such as Naive Bayes, C4.5, Linear Regression, and Least Median Square. This approach can help farmers decide whether the soil is suitable for a particular crop.[13]

The Crop Selection Method (CSM) was proposed in this research paper by Rakesh Kumar to solve the crop selection problem and optimize the net yield rate of crops over the season, resulting in maximum economic growth for the country. The proposed strategy has the potential to improve net crop production rates. [14]

 DOI : 10.35291/2454-9150.2022.0092 

ISSN : 2454-9150 Vol-07, Issue-12, MAR 2022

# IV. EXPERIMENTAL SETUP

The following is a step-by-step approach for designing the crop recommendation system: To begin with, prepare the input dataset in a comma-separated values file that contains the crop recommendation dataset that must be pre-processed. The input dataset is then put through a variety of pre-processing procedures, including missing value filling, categorical data encoding, and scaling of values within a reasonable range. Based on the selected split ratio, the pre-processed dataset is separated into training and testing datasets. The suggested study uses an 80:20 split ratio, which indicates that 80 percent of the dataset is used to train the model and the remaining 20% is used as a test dataset. After that, Naive Bayes, Decision Tree, XG Boost, SVM classifier, and Random Forest classifier are used to train the model.

# V. PROPOSED SYSTEM

In this study, the system proposes an Intelligent Crop Recommendation system that predicts crop suitability by taking into account all relevant parameters such as temperature, rainfall, location, and soil condition. In addition, provide organic fertilizer recommendations based on what soil nutrients the soil lacks or has an abundance of. The project is divided into two subsystems: one that recommends crops and the other that recommends organic ways to improve soil quality for growing a specific crop.In Fig.1 the block diagram represents the flow of steps in developing the machine learning model.

## A. DATA ANALYSIS

The number of parameters and the correctness of the training dataset determines the accuracy of any machine learning algorithm. The system uses a custom-built dataset for the first sub-system by combining various datasets from government websites and Kaggle datasets. This dataset is created by augmenting existing rainfall, climate, and fertilizer datasets for India.

## B . DATA PREPROCESSING

There are two steps to this process. The first step is to remove the missing values in the original dataset, which were represented by a dot ('.'). The presence of missing values deteriorates the value of the data and, as a result, the performance of machine learning models is hampered. Therefore, to deal with these missing values, it is replaced with large negative values that the trained model can easily identify as outliers. Before the data can be used by machine learning algorithms, the second step is to generate class labels. Class labels are required because it is intended to use supervised learning. Because the original dataset lacked labels, we had to create them during the data preprocessing phase.

## C. MACHINE LEARNING ALGORITHMS

After the preprocessing step, the data-set is used to train different machine learning models using algorithms like Decision Tree, K Nearest Neighbor (KNN), Random Forest Classifier, SVM classifier, XG boost are five machine learning algorithms that have in-built support for MLC.

### DECISION TREE

It is a supervised learning algorithm that uses a tree to represent attributes and class labels. Here, root attributes are compared to record attributes, and a new node is reached based on the result of the comparison. This process is repeated until a leaf node with a predicted class value is found. As a result, a modeled decision tree is very effective for prediction. Decision trees are a highly effective structure for laying out possibilities and investigating the potential implications of selecting those options.

### LOGISTIC REGRESSION

Under the Supervised Learning approach, logistic regression is precise in outputs for various use cases. It's a technique for predicting a categorical dependent variable from a collection of independent variables. A categorical dependent variable's output is predicted using logistic regression. Therefore, the result must be a discrete or a categorical value. The value can be Yes or No, 0 or 1, true or false, etc but instead of giving exact values like 0 and 1, it generates probabilistic values that are somewhere between 0 and 1. The classification problems are solved using logistic regression.

### SUPPORT VECTOR MACHINE (SVM)

The "Support Vector Machine" (SVM) is a supervised machine learning technique that can solve classification and regression problems. It is, however, mostly employed to solve categorization problems. Each data item is plotted as a point in n-dimensional space (where n is the number of features you have), with the value of each feature being the value of a certain coordinate in the SVM algorithm. Then we accomplish classification by locating the hyper-plane that clearly distinguishes the two classes. SVMs, in their most basic form, find a separation line (or hyperplane) between data from two classes. SVM is an algorithm that takes data as input and, if possible, generates a line that separates the classes.

### RANDOM FOREST

Random Forest is a well-known machine learning algorithm from the supervised learning technique. It can be applied to both classification and regression issues in machine learning. It is built on the notion of ensemble learning, which is the process of merging numerous classifiers to solve a complex problem and improve the model's performance. "Random Forest is a classifier that contains several decision trees on various subsets of the provided dataset and takes the average to increase the predicted

 DOI : 10.35291/2454-9150.2022.0092 

accuracy of that dataset," as the name implies. Instead of relying on a single decision tree, the random forest collects the output from each tree and predicts the final output based on the majority votes of predictions. The greater the number of trees in the forest, the higher the accuracy and the lower the risk of overfitting. [8] . In Fig 2 the algorithm for random forest is depicted.

NAÏVE BAYES

Naive Bayes is a classification algorithm for binary and multi-class classification problems. When binary or categorically input values are provided, the Naive Bayes method is very easy to understand. In Naive Bayes, a Naive Bayes classifier assumes that the presence of a particular feature in a class is not at all related to the presence of any other feature. Hence the name 'Naive Bayes'.

D. BLOCK DIAGRAM

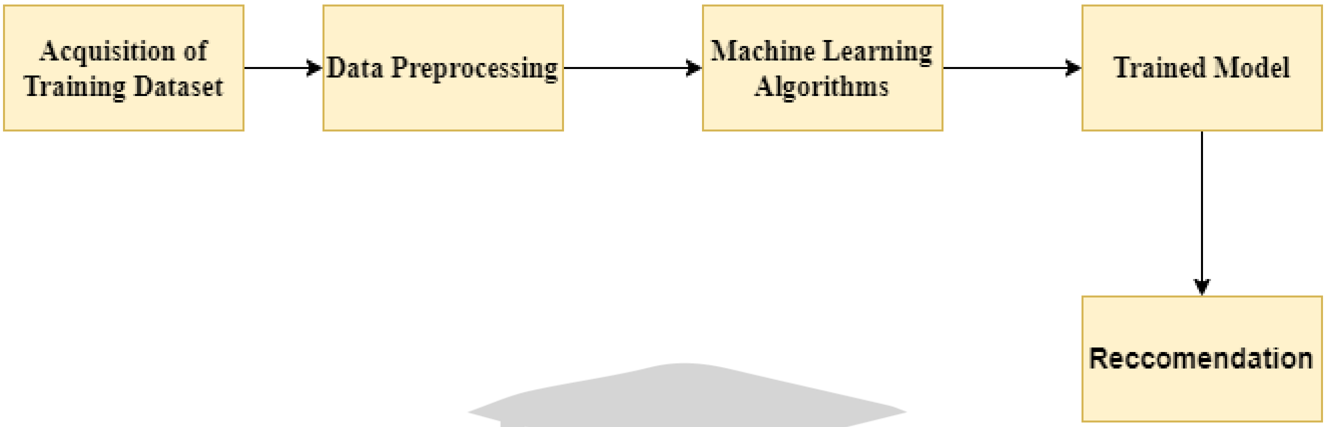


Fig.1: Crop Recommendation System

E. ALGORITHM (PSUEDOCODE)

**Algorithm 1: Pseudo code for the random forest algorithm**

To generate $c$ classifiers:
**for** $i = 1$ to $c$ **do**
Randomly sample the training data $D$ with replacement to produce $D_i$
Create a root node, $N_i$ containing $D_i$
Call BuildTree( $N_i$ )
**end for**

**BuildTree(N):**
**if** $N$ contains instances of only one class **then**
**return**
**else**
Randomly select x% of the possible splitting features in $N$
Select the feature $F$ with the highest information gain to split on
Create f child nodes of $N$ , $N_1,\ldots, N_f$ , where $F$ has $f$ possible values ( $F_1, \ldots, F_f$ )
**for** $i = 1$ to $f$ **do**
Set the contents of $N_i$ to $D_i$ , where $D_i$ is all instances in $N$ that match $F_i$
Call BuildTree( $N_i$ )
**end for**
**end if**

**Fig.2 Algorithm for Random Forest**

 DOI : 10.35291/2454-9150.2022.0092 

F. RESULTS

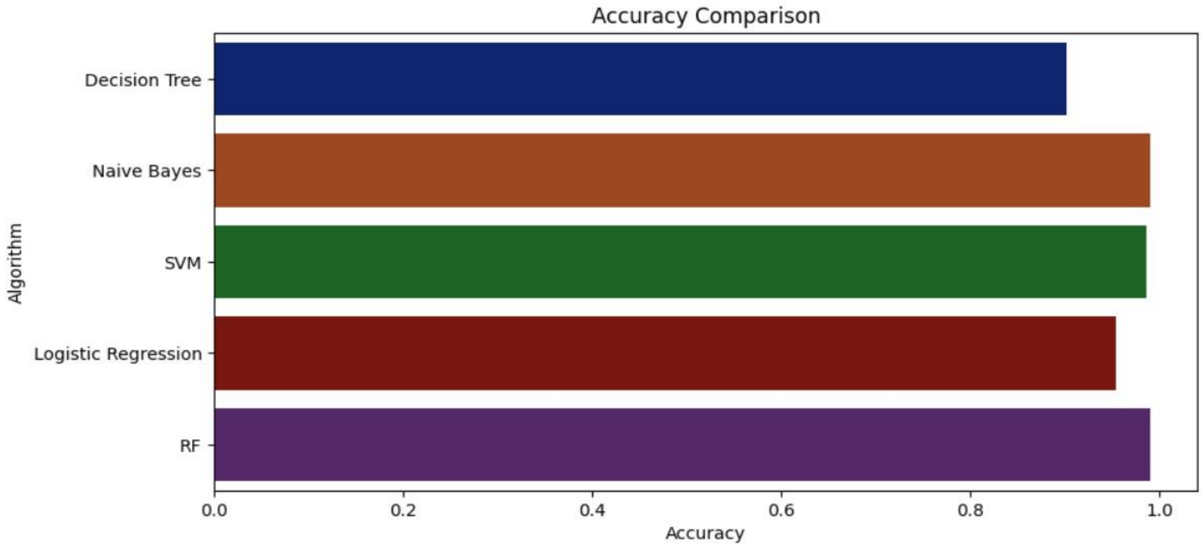


Fig.3 : Accuracy Comparison

**Making a prediction**

```
In [37]:  data = np.array([[104,18, 30, 23.603016, 60.3, 6.7, 140.91]])
          prediction = RF.predict(data)
          print(prediction)

          ['coffee']

In [38]:  data = np.array([[83, 45, 60, 28, 70.3, 7.0, 150.9]])
          prediction = RF.predict(data)
          print(prediction)

          ['jute']

In [39]:  data = np.array([[90, 42, 43, 20.8, 82, 6.57, 202.935]])
          prediction = RF.predict(data)
          print(prediction)

          ['rice']

In [40]:  data = np.array([[43, 68, 81,17.47,17.932, 6.76,78.9206]])
          prediction = RF.predict(data)
          print(prediction)

          ['chickpea']

In [ ]:
```

**Fig.4 Crop Prediction Result**

## VI. CONCLUSION

The proposed system is implemented along with the algorithm namely, Random Forest Classifier using Pandas, Numpy, Tensorflow, seaborn and Scikit-learn libraries, tools, and Python as the programming language. The Random forest classifier model is selected because it has higher accuracy compared to all the other models. Fig.3 shows a comparison of accuracy between different algorithms. Fig.4 represents the final outcome or prediction result of the trained model. The suggested approach assists farmers in selecting the best crop by delivering information that conventional farmers do not keep track of lowering the likelihood of crop failure and boosting productivity. It also stops them from sustaining losses. It is planned for the future to add a chatbot in the regional language to provide agricultural recommendations to the farmers that can be accessed by millions of farmers across the country.

## REFERENCES


[1] https://censusindia.gov.in/census_and_you/economic_activity.aspx#. *(references)*

[2] S.Pudumalar, E.Ramanujam, "Crop Recommendation System for Precision Agriculture", 2016, IEEE Eighth International Conference on Advanced Computing (ICAC)

[3] M. T. Shakoor, K. Rahman, S. N. Rayta and A. Chakrabarty, "Agricultural production output prediction using Supervised Machine Learning techniques", 1st International Conference on Next Generation Computing Applications (NextComp), Mauritius, 2017, pp. 182-187, DOI: 10.1109/NEXTCOMP.2017.8016196.

[4] 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) | 978-1-6654-0360-3/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICCMC51019.2021.941837


 DOI : 10.35291/2454-9150.2022.0092 

[5] "Agriculture in India: Industry Overview, Market Size, Role in Development...|IBEF", available at https://www.ibef.org/industry/agriculture-india.aspx, visited in February 2018

[6] 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEAM. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] Priyadharshini A, Swapneel Chakraborty, Aayush Kumar, Omen Rajendra Pooniwala "Intelligent Crop Recommendation System using Machine Learning",2021 5th International Conference on Computing Methodologies and Communication (ICCMC) | 978-1-6654-0360-3/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICCMC51019.2021.9418375

[8] https://www.javatpoint.com/machine-learning-random-forest-algorithm

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

[10] T.R. Lekhaa, "Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques", International Journal of Modern Trends in Engineering and Science (IJMTES), 2016, Volume 03, Issue 10

[11] Nidhi H Kulkarni, Dr. G N Srinivasan Dr. B M Sagar Dr.N K Cauvery, "Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique"3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018 ISBN: 978-1-5386-6078-2 © 2018 IEEE.

[12] Lokesh.K, Shakti.J, Sneha Wilson, Tharini. M.S, "Automated crop prediction based on efficient soil nutrient estimation using a sensor network", July 2016, National Conference on Product Design (NCPD 2016)

[13] 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

[14] Crop Selection Method to Maximize Crop Yield Rate using Machine Learn ing Technique Authors: Rakesh Kumar, M.P. Singh, Prabhat Kumar, and J.P. Singh

 DOI : 10.35291/2454-9150.2022.0092 