

Implementation of Crop Selection by Land Dataset Using Machine Learning

Gayathri U , Dr.Priya Anand

Vels Institute of Science , Technology and Advanced Studies, India

DOI : <https://doi.org/10.51583/IJLTEMAS.2025.140400049>

Received: 28 April 2025; Accepted: 01 May 2025; Published: 09 May 2025

Abstract: As we are aware of the fact that, most of Indians have agriculture as their occupation. Farmers usually have the mind-set of planting the same crop, using more fertilizers and following the public choice. By looking at the past few years, there have been significant developments in how machine learning can be used in various industries and research. So, we have planned to create a system where machine learning can be used in agriculture for the betterment of farmers. India is an Agricultural Country and its economy largely based upon crop productivity. So we can say that agriculture can be pillar of all business in our country. Selecting of always crop is very important in the agriculture planning. Many researchers studied guess of yield rate of crop, guess of weather, soil categorizing and crop classification for agriculture planning using machine learning techniques.

Many changes are required in the agriculture department to improve changes in our Indian economy. We can improve agriculture by using machine learning system which are applied simply on farming sector. Along with all advances in the machines and technologies used in farming, functional information about different matters also plays a significant role in it. The concept of this paper is to implement the crop selection method so that this method helps in solving many agriculture problems. This enhances our Indian wealth by maximizing the yield rate of crop production. In our project crop is predicted by algorithm namely Recurrent Neural Network (RNN) as proposed and Random Forest (RF) as existing and its accuracy is calculated and compared with other algorithms.

I. Introduction:

Agriculture plays a very important role where economic growth of a country like India is considered. The main aim of agricultural planning is to achieve maximum yield rate of crops by using little number of land resources. Many machine learning algorithms can help in raising the producing of crop yield rate. Whenever there is loss in critical situation, we can apply crop selecting method and reduce the Losses. And it can be used to acquire crop yield rate in favorable conditions. This Maximizing of yield rate helps in upgrade countries economy. We have some other factors that impact the crop yield rate. They are fertilizer quality and crop Selection. Selection of crops depends upon two task that are favorable and critical conditions. Many researches are carried out to improve agricultural setting. The aim is to get the biggest yield of crops. Many categorizing systems are also applied to get maximum yield of crops. Machine learning can be used to better the yield rate of crops. The method of crop Selection is bother to improve crop production.

The construction of crops may depend on geographical conditions of the region like river ground, hill areas or the depth areas. Weather conditions like humidity, rainfall, temperature, cloud. Soil type may be clay, sandy, salt or peaty. Soil composition can be copper, potassium, sulphate, nitrogen, manganese, iron, calcium, ph. value or carbon and different methods of harvesting. Many parameters are used for different Crops to do separate predictions. These prediction models can be studied by using analyzer. These predictions are classified as two types. One is traditional statistic method and other is machine teaching techniques. Traditional method helps in predicting single sample spaces. And machine learning methods helps in predicting different predictions. We need not to consider the design of data models in traditional method where as we need to consider the structure of data models in machine learning methods.

Existing System:

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance.

Decision trees are a popular method for various machine learning tasks. Tree learning "come[s] closest to meeting the requirements for serving as an off-the-shelf procedure for data mining", say Hastie et al., "because it is invariant under scaling and various other transformations of feature values, is robust to inclusion of irrelevant features, and produces inspectable models. However, they are seldom accurate".

In particular, trees that are grown very deep tend to learn highly irregular patterns: they overfit their training sets, i.e. have low bias, but very high variance. Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same

training set, with the goal of reducing the variance. This comes at the expense of a small increase in the bias and some loss of interpretability, but generally greatly boosts the performance in the final model.

Forests are like the pulling together of decision tree algorithm efforts. Taking the teamwork of many trees thus improving the performance of a single random tree. Though not quite similar, forests give the effects of a K-fold cross validation.

Purpose

Smart farming using machine learning and data analytics' is a POC intended for betterment of farmers and agriculture in India. Mentioned in the report are the exact goals to be fulfilled by this project – namely, recommendation of the top crops based on soil properties and environment factors. These take into consideration all the nutrients and micronutrients with reference to the soil health card. From these top crops recommended, the farmer will also be able to know the economically beneficial crop where the model implements price prediction module for the crops.

Also a year round plan will be provided to farmer to maximize his cultivation and earnings. Besides generic information regarding government schemes or soil testing labs will be provided. The researcher along with these functionalities can access certain datasets that he wants to explore. The researcher can also study the impact of various algorithms on varies data. An add on 'what if' scenario will help not only the farmer but also there searcher to play around with various factors and see the variations in results and proceed with decision making accordingly.

Implementation

Implementation of Crop Selection by Land Dataset Using Machine Learning" aims to recommend the most suitable crop for cultivation based on land and climatic conditions. Initially, a relevant dataset containing features such as soil type, soil pH, temperature, rainfall, and humidity is collected. The data undergoes preprocessing steps, including handling missing values, encoding categorical variables, and scaling numerical features to ensure consistency. After preprocessing, feature selection techniques are applied to identify the most influential parameters impacting crop yield.

For the machine learning model, a classification approach is adopted, with algorithms such as Random Forest, Decision Tree, and Support Vector Machine considered. The dataset is split into training and testing sets to evaluate model performance. The Random Forest Classifier, due to its robustness and accuracy, is selected as the primary model. After training, the model's performance is assessed using accuracy score, confusion matrix, and classification report metrics.

Upon achieving satisfactory results, the model is saved for deployment purposes using serialization techniques like joblib. An optional web-based application can be developed using frameworks like Flask, allowing users to input land parameters and receive crop recommendations in real time. This project not only streamlines the crop selection process for farmers but also promotes optimal land utilization based on scientific data analysis. Future enhancements could involve integrating real-time weather data, fertilizer recommendations, and improving model accuracy with advanced ensemble methods.

II. Conclusion

After going through many surveys and through analysis we can conclude that the use of no of various machine learning algorithms will not only help farmers to get better results but also increase their revenue which for many is a matter of life and death. Currently farmers make rough estimations based on their previous experiences and plan accordingly, using ML instead will definitely decrease the margin of error and provide them with better outputs. This proposed system will work to provide suggestions, which definitely going to helps farmers to get more yield and better crops.

Future Enhancement

We will do research on following area of recommendation as well as pricing. We will try to consider both user and providers concerns of changing demand and its cost. This will ensure both provider and customers benefit. Apart from this we will consider competitive prices and its result on pricing. We will study best fit auction based pricing to support optimized fine grained scheme. Also partial waste issue is a area of study which can result in reduced prices using precise scheduling of users' job. User scheduling behaviors and partial usage waste will be brainstormed to find an effective solution.

Reference

1. Mayank Champaneri, ChaitanyaChandvidkar, DarpanChachpara, MansingRathod, "Crop yield prediction using machine learning" International Journal of Science and Research, April 2020.
2. Pavan Patil, VirendraPanpatil, Prof.ShrikantKokate, "Crop Prediction System using Machine Learning Algorithms", International Research Journal of Engineering and Technology, Feb 2020.
3. Ramesh Medar, Shweta, Vijay S. Rajpurohit, "Crop Yield Prediction using Machine Learning Techniques", 5th International Conference for Convergence in Technology, 2019.
4. TruptiBhange, Swati Shekapure, Komal Pawar, HarshadaChoudhari, "Survey Paper on Prediction of Crop yield and Suitable Crop", International Journal of Innovative Research in Science, Engineering and Technology, May 2019.

5. E. Manjula, S. Djodiltachoumy, "A Modal for Prediction of Crop Yield", International Journal of Computational Intelligence and Informatics, March 2017.
6. Nishit Jain, Amit Kumar, SahilGarud, Vishal Pradhan, Prajakta Kulkarni, "Crop Selection Method Based on Various Environmental Factors Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Feb 2017.
7. Rakesh Kumar, M.P. Singh, Prabhat Kumar, J.P. Singh, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, T.N., India., May 2015.
8. Liu Qingyuan, and Wang Tianchuan, "Market price theory and practice". Dalian: Northeast University of Finance and Economics Press, 1998, pp. 15-20.
9. R. S. Pindyck, and D. L. Rubinfeld, "Econometric Models and Economic Forecasts", The McGraw-Hill Companies, Inc, 1998, pp. 251-260.
10. Yao Xia, Peng Hangen, and Zhu Yan, "ARIMA Time Series Modeling and Applying on Fresh Agricultural Products", System Sciences and Comprehensive Studies in Agriculture, vol. 23, Feb. 2007 , pp. 88-94.
11. Nie Rong, Qin Keming, and Zhang Xiaohong, "Stochastic Model and Risk Measuring on the Farm-produce's Price", Mathematics In Practice and Theory, vol. 34, Nov. 2004, pp. 108-112.
12. Zhang Xiaoshuan, "Research of Aquatic Products Price Forecasts Support System", Beijing: China Agricultural University, 2003, pp. 10-15.
13. Yu Shouhua, Huang Haoran, and Ou Jingying, "Vegetables Prices Prediction Research of Regional Agricultural Products Wholesale Market", Research of Agricultural Modernization, 27 monog, Dec. 2006, pp. 118-120.
14. Cheng Xianlu, "Vegetables Prices Forecast and Predict System Research of Beijing's Agricultural Products Wholesale Market", Beijing Agricultural Sciences, Feb. 2002, pp. 1-10.
15. Yuan Zenren, "Artificial neural network and applications", Beijing: Qinghua University Press, 1998, pp. 13-14.
16. MT Hagan, and MB Menhaj, "Training Feed Forward Networks with Marquart Algorithm", IEEE Trans.on Neural Networks, vol. 5, June 1994, pp. 989-993.
17. Xu Dong, and Wu Zhen, "System analysis and design-neural network based on the MATLAB6.X" (the 2nd version), Xi'an: Xi'an Electronic Technology University Press, 2002, pp. 30-33.
18. Li Xiaofeng, "New improvement of BP neural network and its application", Journal of Jilin Institute of Chemical Technology, vol. 17, Dec. 2000, pp. 48-51.
19. Gao Ling, "The crop pest situation forecasting and achieving in MATLAB based on BP neural network", Hefei: Anhui Agricultural University, 2003, pp. 45-52.
20. Jiang Shaofei, "The method of data processing of artificial neural networks in civil engineering problems", Journal of Harbin University of Civil Engineering and Architecture, vol. 32, Oct. 1999, pp. 24-28.
21. Liu Yongjian, and Liu Yijian, "The improving methods of artificial neural network in Geo-technical engineering", Journal of Guangdong University of Technology, vol. 19, Mar. 2002, pp. 21-25.
22. Liu Wenxi, "New Keynesian price rigidity micro-theoretical foundation", Economic Science, May 1997, pp. 60-8