Proceedings of the SMART–2021, IEEE Conference ID: 52563
10th International Conference on System Modeling & Advancement in Research Trends, 10th–11th December, 2021
Faculty of Engineering & Computing Sciences, Teerthanker Mahaveer University, Moradabad, India

2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART) | 978-1-6654-3970-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMART52563.2021.9676210

Crop Recommendation System using Machine Learning Algorithms

Gaurav Chauhan[1], Alka Chaudhary[2]
[1,2]*Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh*
E-mail: [1]gauravch13004@gmail.com, [2]achaudhary4@amity.edu

***Abstract*—India is a predominantly agricultural country, with agriculture playing animportant part in the Indian economy and people's lives. Crops are recommended based on soil, weather, humidity, rainfall, and other variables to increase agricultural output. It benefits not just farmers, but also the country and helps to keep food costs down.**

This paper presents the utilisation of machine learning approaches like Random Forest and Decision Tree to predict which crop is best for which soil type based on the data sets.
Keyword: Machine Learning Techniques, Recommendation System.

I. Introduction

Agriculture is one of India's most important activitity. It is the most diverse economic sector and the most significant contributor to world progress. Previous crop forecasting and crop predicting are based on farmer experience in a specific location. They will prefer an earlier or neighbouring crop, or a trendy crop in the surrounding region, solely in their area, and lack appropriate knowledge of the soil's nutrient content. Today, thanks to technological advancements, a variety of technologies are available to handle the requirement for agricultural inputs. This paper looks at a few useful systems that make use of machine learning techniques.

II. Machine Learning Algorithms

A. Linear Regression

By introducing a line queue into the data, line deceleration tries to represent the relationship between the two variables. Some variables are referred to as descriptive variables, while others are referred to as dependent variables. A regression line model, for example, might be used to connect individual weights with their length.[1][2].

The a and b coefficients are obtained by minimising the sum of the squared differences in the distance between the data points and the regression line[1][2].

B. Logistic Regression

Subtraction is used to estimate distinct values from a set of independent variables (typically binary values equal to 0/1). By adding data into a log activity, it can help anticipate the occurrence of an event. Log drop is another name for it.

To enhance logistic regression models, the following approaches are often used:

- Use a nonlinear model
- Add interaction terms
- Remove functionality
- Control methods

C. Decision Tree

One of the most often used machine learning methods nowadays is supervised learning, which is used to categorise situations. It's suitable for both phase-based and continuous applications. Based on the most essential / different independent criteria, this strategy separates the population into two or more identical groups.

D. SVM (Doctor of Science in Veterinary Medicine)

SVM is a classification algorithm that uses n-dimensional points to represent raw data. Each element's value is subsequently connected to a specific integration, making data sharing more easier.Lines that may be used to segregate data and show it on a graph are known as classifiers.

E. Naive Bayes

The Bayesian naive classifier assumes that one element in one category exists independently of the existence of another member in another category.

Despite the fact that these qualities are linked, a Naive Bayesian classifier would look at each one independently when calculating the likelihood of a specific outcome.

A naïve Bayesian model is easy to build and can handle large data sets. It's straightforward, and we've demonstrated that it beats even the most complex categorization algorithms.

F. KNN (K- Nearest Neighbors)

Separation and retreat issues can be addressed with this strategy. It appears to be widely utilised in the data science business to tackle categorization challenges. It's a basic algorithm that recalls all prior examples and distinguishes all new scenarios based on a majority decision made by its neighbours. The case is then assigned to a class that is very similar to it.

G. K-means

To deal with the intricacy of integration, this unattended method is adopted. Data sets are separated into a number of clusters (let's call them K) so that each cluster's data points are distinct and distinct from the data in other clusters.

- Each data point generates a cluster with the focus points nearest to it, resulting in k clusters, the K-Means method, which identifies k points, known as centroids, for each cluster.
- New focus points have been developed based on the current cluster members.

III. Literature Review

A. Agriculture is India's primary source of revenue and employment. The most common issue that Indian farmers encounter is that they do not choose the right crop for their land. As a result, they will see a considerable reduction in production. Precision agriculture has been utilised to help farmers overcome their problems. Precision agriculture is a contemporary agricultural approach thuses research data on soil characteristics, soil types, and crop production statistics to advise farmers on the optimum crop for their specific site. This reduces the number of times a crop is selected erroneously while also increasing productivity[1].

B. Predicting crop yields well ahead of harvest can help farmers and government agencies make informed decisions about how to store, sell, set minimum support prices, import/export, and other issues. Predicting a crop requires a thorough study of huge quantities of data generated from a variety of factors such as soil quality, pH, EC, N, P, K, and so on. This prediction approach is suitable for data mining since crop prediction requires a large number of databases. Data mining is a technique for extracting knowledge from large quantities of data. The numerous data mining methods that have been used to estimate agricultural yields are discussed in this document. Any crop production prediction system's success is determined on the accuracy with which characteristics are extracted and how well classifiers are employed. This paper summarises the results of a variety of agricultural output forecast algorithms used by various authors, as well as their accuracy and suggestions [2].

C. In this paper, we present RSF, a farmer recommendation system that can propose the best crops to produce in diverse locations. The system recognises a user's location before calculating similarity across upazilas utilising different agro-ecological and agro-climatic data at the upazila level using the Pearson co-relation similarity technique. The top-n upazilas that are comparable are then picked. Finally, it recommends top-fc crops to an upazila user based on seasonal data and crop production rates in similar upazilas. We put the system to the test with real data and observed that it was accurate enough. Farmers may use the approach to help them cultivate the correct crops.As a result, they will be able to enhance their quality of life and make greater contributions to society. We made the system layout available in Bangla and English so that farmers and anyone who work with them may easily understand it[3].

D. Because numerous factors impact the entire agricultural process, the decision-making process is highly difficult. This app provides a list of the most likely crops for that farm as a result of a backend computation. The suggested decision support system may assist farmers in selecting a crop for cultivation mapping by taking into account a variety of factors such as soil type, PH value, average weather requirements, needed water consumption, temperature range, and so on[4].

E. Agricultural intelligent decision systems are useful for guiding agricultural output and can offer a scientific foundation for agricultural research. Intelligent decision systems can benefit from the usage of big data analysis technologies. The research and development of agricultural intelligent decision systems is explored. For the first time, the agricultural decision system is categorised. The frame designation of the intelligent decision system is examined, and the design method is presented [5].

F. By analysing an agriculture dataset, this project focuses on applying Data Mining techniques to develop a crop production forecast system. Different classifiers used for prediction, and their performance is compared using the WEKA tool. The technique will be more accurate if the error value is smaller. A comparison of classifiers yielded the result [6].

G. Agriculture is highly important in India. Farmers flourish, and the country prospers as a result. Farmers will be able to sow the proper seed based on soil conditions as a consequence of our efforts, improving the country's output. Our future efforts will be focused on improving the data set by adding more characteristics and including yield prediction [7].

H. Agriculture is our country's primary source of revenue and the economic backbone. Farmers must be prepared to deal with contemporary challenges such as water shortages, uncontrolled costs due to demand-supply mismatches, and weather instability by practising smart farming. Low agricultural yields as a result of unforeseen climate changes, insufficient irrigation infrastructure, soil fertility deterioration, and conventional farming methods must all be addressed. Machine learning is one such technique used in agriculture to forecast crop production. Many machine learning techniques, such as prediction, classification, regression, and clustering, are utilised to anticipate agricultural output [8].

I. This research examines the requirements and planning required to construct a precision farming software model. It is compatible with the appropriate agricultural framework. The writers start with the fundamentals of precision farming before moving on to support development and modelling. Using the Precision Agriculture (PA) principle, this study

presents a technique for managing diversity on small, open farms for each farmer and crop quality. The model's major goal is to provide detailed guidance to even the tiniest farmer at the crop clip level. SMS and email are two of the most extensively utilised technologies. This model was developed for Kerala, which has the smallest average catch size in India. As a result, this mode is now available [9].

J. Agriculture is the most vital field on the planet, and it is the backbone of the Republic of India. For a long time, agriculture has been in chaos. The bulk of agricultural crops are overripe in terms of productivity due to the influence of temperature variations and accompanying uncertainties. an appropriateCrop expansion forecasting is a crucial part of crop forecasting. The capacity of the federation industries to provide for its workers will be hampered by such projections. Machine learning is a method for discovering new models from large data sets. Random forest and linear regression are two examples of regressive tec hniques.

Area and production are two types of meteorological data generated by required data. The agricultural yield recommendation is determined in this study by comparing multiple machine learning ML regressions and calculating the total percentage improvement over numerous years [10].

K. Agriculture and associated industries are, without a question, rural India's most vital sources of income. Agriculture makes a substantial contribution to the country's Gross Domestic Product (GDP). The agriculture industry's huge size is a blessing to the country. However, when compared to global norms, agricultural output per acre is low. This might be one of the causes for India's rising marginal farmer suicide rate. This study provides farmers with a realistic and user-friendly production forecast technique [11].

IV. Comparision of Different Research Paper on Crop Recomondation Using Machine Learning

1.	S.Pudumalar, E.Ramanuja m	2016	Crop Recommendation System for Precision Agriculture[1]	Random tree, KNN, Naïve Bayes, WEKA tool	1. Pre-processing of data 2. Extraction of features 3. Detailed model for more accuracy 4. Missing and out-of-range values management 5. Formulation of Rules
2.	Yogesh Gandge, Sandhya	2017	A Study on Various Data Mining Techniques for Crop Yield Prediction[2]	Multiple Linear Regression Decision Tree using ID3, Attribute selection, SVM Neural Networks K-means and KNN, C4.5	1. Choose an agricultural field 2. Choosing Before you plant, harvest. 3. The user's receipt 4. Pretreatment 5. Characteristic selection
3.	Miftahul Jannat Mokarrama	2017	RSF: A Recommendation System for Farmers[3]	Location Detection, Data analysis and storage, Similar location detection, Recommendation generation module.	1. Thermographer, physiographer, Cultureepoch, cultural, Cultivation efficiency 2. It is seasonal Database of Plants 3. Identical location recognition 4. Produce each and every to gather 5. Identicality Between cultivated plants in a location
4.	Prof. Rakesh Shirsath	2017	Agriculture decision support system using data mining[4]	Subscription based system ,Android application, ANN, Personalized content	1. A connection module in an Android application 2. Prior to cultivated crops well-known for system 3. The user Feedback mechanism 4. maintenance Harvest.
5.	Ji-chun Zhao, Jian-xin Guo	2018	Big Data Analysis Technology Application in Agricultural Intelligence Decision System[5]	Inference engine, expertise, Knowledge engineering, Knowledge base for recommendation system	1. A fantastic database of harvestables 2. Make a contract Hadoop is a big data platform. 3. Professional Knowledge 4. The past experience 5. Purpose Selection HDFS (High-Definition File System) 6. Future Applicability: Hadoop with Hadoop Artificially Neural Networks.
6.	Shruti Mishra Priyanka Paygude	2018	Use of Data Mining in Crop Yield Prediction[6]	LWL J48, LAD tree, IBK algorithm	1. Use the WEKA Tool 2. The Following Tree exhibited the cheapest precision 3. Errors can occur be kept to a minimum by chopping the oak tree 4. IBK stood for noticed at Afford higher precision
7.	D. Anantha Reddy, Bhagyashri Dadore, Aarti Watekar	2019	Crop Recommendation System to Maximize Crop Yield in Ramtekregion[7]	1. Random Tree 2. K-Nearest Neighbor 3. Random Forest 4. Decision Tree	1.Agriculture with precision 2.Assembly of the model 3.Voting methods used by the majority Chaid 4. Nearest Neighbo 5.Naive Bayes (Naive Bayes)

8.	Kodimalar Palanivel	2019	An Approach For Prediction Of Crop Yield Using Machine Learning And Big Data Techniques[8]	1. Linear Regression 2. Artificial neural Networks 3. Support vector Machine	1. restoring the image 2. the fidelity of l2 data phrase 3.function of regularisation 4. variation in total
9.	Rohit Kumar Rajak, Ankit Pawar, MitaleePendke	2017	Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique[9]	1. Support Vector Machine 2. NAÏVE Bayes 3. Multi-layer Perceptron 4. Random Forest	1.Agriculture with precision 2.ensemble model 3.recommendation system 4.SVM (Systematic Variable Model 5.ANN Random Tree 6. NBclassifier is number seven.
10.	Alok Kumar Jagadev	2021	Agricultural Recommendation System for Crops Using Different Machine Learning Regression Methods[10]	1. Linear Regression Prediction, Machine Learning, 2. Polynomial Regression 3. Random Forest Regression, 4. Support Vector Regression	1. Accuracy 94.78% 2. The majority voting method has been applied
11.	Shilpa Mangesh Pande; Prem Kumar Ramesh; Anmol Anmol; B. R Aishwarya	2021	Crop Recommender System Using Machine Learning Approach[11][2][13] [14]15][16]	1. Support Vector Machine 2. Artificial Neural Network 3. Random Forest 4. Multivariate Linear Regression 5. K-Nearest Neighbour	1. Accuracy 95% 2. It suggests the best time to use the fertilizers to boost up the yield. 3. The Random Forest showed the best results

V. Conclusions

Smallholder farmers in India have a high suicide rate, according to the country's high suicide rate. This study provides armers with a simple and practical technique to anticipate harvests. Farmers are connected to the planned system via a smartphone application. The user can be located with the use of GPS.

References

[1] Pudumalar, S., et al. "Crop recommendation system for precision agriculture." *2016 Eighth International Conference on Advanced Computing (ICoAC)*. IEEE, 2017.

[2] Gandge, Yogesh. "A study on various data mining techniques for crop yield prediction." *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*. IEEE, 2017.

[3] Mokarrama, Miftahul Jannat, and Mohammad Shamsul Arefin. "RSF: A recommendation system for farmers." *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 2017.

[4] Shirsath, Rakesh, et al. "Agriculture decision support system using data mining." *2017 International Conference on Intelligent Computing and Control (I2C2)*. IEEE, 2017.

[5] Zhao, Ji-chun, and Jian-xin Guo. "Big data analysis technology application in agricultural intelligence decision system." *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*. IEEE, 2018.

[6] Mishra, Shruti, et al. "Use of data mining in crop yield prediction." *2018 2nd International Conference on Inventive Systems and Control (ICISC)*. IEEE, 2018.

[7] Reddy, D. Anantha, BhagyashriDadore, and Aarti Watekar. "Crop recommendation system to maximize crop yield in ramtek region using machine learning." *International Journal of Scientific Research in Science and Technology* 6.1 (2019): 485-489.

[8] Palanivel, Kodimalar, and ChellammalSurianarayanan. "An approach for prediction of crop yield using machine learning and big datatechniques." *International Journal of Computer Engineering and Technology* 10.3 (2019): 110-118.

[9] Rajak, Rohit Kumar, et al. "Crop recommendation system to maximize crop yield using machine learning technique." *International Research Journal of Engineering and Technology* 4.12 (2017): 950-953.

[10] Garanayak, Mamata, et al. "Agricultural Recommendation System for Crops Using Different Machine Learning Regression Methods." *International Journal of Agricultural and Environmental Information Systems (IJAEIS)* 12.1 (2021): 1-20.

[11] PANDE, SHILPA MANGESH, et al. "Crop Recommender System Using Machine Learning Approach." *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2021.

[12] Chaudhary, Alka, V. N. Tiwari, and Anil Kumar. "Design an anomaly based fuzzy intrusion detection system for packet dropping attack in mobile ad hoc networks." *2014 IEEE International Advance Computing Conference (IACC). IEEE,* 2014.

[13] Chaudhary, Alka, V. N. Tiwari, and Anil Kumar. "A new intrusion detection system based on soft computing techniques using neuro-fuzzy classifier for packet dropping attack in manets." *International Journal of Network Security 18.3 (2016): 514-522.*

[14] Chaudhary, Alka, V. N. Tiwari, and Anil Kumar. "A cooperative intrusion detection system for sleep deprivation attack using neuro-fuzzy classifier in mobile ad hoc networks." *Computational Intelligence in Data Mining-Volume 2. Springer, New Delhi, 2015. 345-353.*

[15] Chaudhary, Alka, Anil Kumar, and V. N. Tiwari. "A reliable solution against packet dropping attack due to malicious nodes using fuzzy logic in MANETs." *2014 International Conference on Reliability Optimization and Information Technology (ICROIT). IEEE, 2014.*

[16] Chaudhary, Alka. "Mamdani and sugeno fuzzy inference systems' comparison for detection of packet dropping attack in mobile ad hoc networks." Emerging technologies in data mining and information security. Springer, Singapore, 2019. 805-811.

