INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD
ISSN(O): 2455-0620 [Impact Factor: 7.581]
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value : 86.87
Volume - 9, Issue - 3, March - 2023 Publication Date: 31/03/2023

DOIs:10.2015/IJIRMF/202303020 --:-- Research Paper / Article / Review

Crop Recommendation System Based On Machine Learning

[1]Dr. M. Deepa, [2]Sowmiya.V, [3]Tamizhan.E, [4]Venkat Vijay.M.P, [5]Sri Ranjani. S.
[1]Associate Professor, [2,3,4,5] Student
[1,2,3,4,5] Department of Computer Science & Engineering, Paavai College of Engineering, Namakkal,
Email - legithasai2010@gmail.com

Abstract: *AWithout a question, the main source of income in rural India is provided by agriculture and its affiliated industries. The country's Gross Domestic Product is also significantly influenced by the agriculture industry (GDP). The country is fortunate to have such a huge agriculture sector. However, the yield of crops per hectare in relation to global norms is unsatisfactory. This is one of the likely reasons why marginal farmers in India commit suicide at a higher rate. For farmers, this research suggests a practical and approachable yield prediction system. A smartphone application used in the proposed method connects farmers to the internet. GPS assists in locating the user. The user enters the location and soil type. The most profitable crop list can be picked using machine learning algorithms, and they can also forecast crop yields for user-selected crops. Selected Machine Learning algorithms, including Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbor (KNN), are used to estimate crop productivity. The Random Forest among them demonstrated the best outcomes with 95% accuracy. The algorithm also makes recommendations on when to apply fertilizers to increase yield. bcdf*

Key Words: *Crop Yield Prediction, Machine Learning, Random Forest, Crop Recommender System, Artificial Neural Networks (ANN), Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Multivariate Linear Regression (MLR), Fertilizer.*

1. INTRODUCTION:

India has a long history of agriculture. India recently gained second place in the world for farm output [15]. Forestry and fishing, two businesses closely allied to agriculture, made up almost 50% of all employment and 16.6% of the nation's GDP in 2009. India's agricultural sector's financial contribution to GDP is declining [1]. The main financial contributor to agriculture is crop yield. Numerous variables, including meteorological, geographic, organic, and financial components, affect crop yield [6]. Due to shifting market prices, farmers find it challenging to decide when and which crops to sow [7]. According to Wikipedia, over the past 10 years, India's suicide rate has increased from 1.4 to 1.8% per 100,000 people [15].

Due to the unpredictability of the weather, farmers are unsure of which crop to plant and when and where to begin. The use of various fertilisers is also unclear because of seasonal climate variations and changes in the availability of fundamental resources like soil, water, and air. The crop yield rate is steadily decreasing in this situation [2]. Giving the farmers access to an intelligent, user-friendly recommender system will solve the issue. In the agricultural industry, crop production prediction is a crucial issue [3]. Every farmer tries to determine the crop yield and whether it matches their expectations [4], valuing their prior knowledge of that particular crop to forecast the output [3]. Weather plays a major role in determining agricultural production.

In this study, we offer a model that takes these concerns into account. The new aspect of the suggested approach is that it gives farmers instructions on how to increase crop yield while also recommending the most lucrative crop for a given area. The suggested approach offers crop selection based on economic and environmental factors, with the goal of maximizing crop yield in order to satisfy the nation's rising food demand [8]. The suggested approach makes crop production predictions by looking at variables like rainfall, temperature, area, season, soil type, etc. The technique aids in choosing the ideal moment to apply fertilizers. The current crop production prediction method is either hardware-based, expensive to maintain, or difficult to use. In this study, we offer a model that takes these concerns into account. The new aspect of the suggested approach is that it gives farmers instructions on how to increase crop yield while also recommending the most lucrative crop for a given area. The suggested approach offers crop selection based on economic and environmental factors, with the goal of maximizing crop yield in order to satisfy the nation's rising food demand [8]. The suggested approach makes crop production predictions by looking at variables like rainfall, temperature, area,

season, soil type, etc. The technique aids in choosing the ideal moment to apply fertilizers. The current crop production prediction method is either hardware-based, expensive to maintain, or difficult to use.

2. LITERATURE REVIEW:

The main efforts done to improve agriculture are integrating technical know-how and innovations to make the agricultural industry more proficient and simple for farmers by anticipating the right crops using all ML techniques. The advantages of several methods, including ANN, Fuzzy Networks, and data mining approaches, are discussed in the paper. A further problem is incorporating real-time datasets with all of them. [9].

One of the earliest efforts created a specific website to evaluate the effect of weather conditions on crop productivity in the designated Madhya Pradesh districts [10]. Based on the area where the crop was grown, the districts were chosen. The first five districts with the largest agricultural area were selected using these criteria. The study's crop selection was based on the dominant crops in the districts that were chosen. The crops chosen were maize, soybean, wheat, and paddy, and the yields for a continuous 20-year period of knowledge were tabulated for each. For the selected crops, the established model's accuracy ranged from 76% to 90%, with an average accuracy of 82%.

Another significant piece of research evaluates the soil's quality and forecasts crop yields together with an appropriate fertiliser suggestion [11]. This model's inputs included the Ph value and the user-provided location. To forecast the weather and temperature for the present location, an API was used. The system compares the outcomes of supervised and unsupervised machine learning methods with accuracy of 82%.

In [12], a classifier that predicts agricultural yield using a greedy technique was put out. It has been demonstrated that using an attribute in a decision tree classifier improves performance. It has been demonstrated that an ensemble model, which is presented, integrates the impacts of several models and is generally superior to the individual models. To forecast agricultural yield, random forests ensemble classification employs many decision tree models. The ratio of 67% to 33% is used to divide the data into two groups, such as training data and test data, from which the mean and standard deviation are computed. To obtain the most accurate findings, our method additionally clusters crops that are related.

In the agriculture industry, a lot of work has been done and several ML algorithms have been used. Increasing agricultural output and providing it to consumers at the greatest price and quality is the largest issue in agriculture. Additionally, it has been noted that at least 50% of farm food is lost and never consumed. The suggested approach offers suggestions for reducing farm product waste. A model that predicts crop yield using KNN algorithms and clustering is presented in one of the more recent papers. KNN clustering, which takes into account the grouping of related crops to provide the most accurate results, has been demonstrated to perform much better than SVM or regression [13].

Enet, Lasso, and Kernel Ridge algorithms, among other sophisticated regression approaches, are used in [17] to forecast the agricultural output for a certain year. The algorithms' precision was improved with the use of stacking regression. Using the Pandas profiling tool, the historical datasets are filtered to extract the datasets for Maharashtra state. The accuracy of the crop yield prediction model was improved by modifying the bias, weight, and Adam optimizer in a multilayer perceptron neural network. The suggested model predicts crop yield using an ANN using a three-layer neural network [18]. Crop production prediction system implementation involves supervised learning. established the relationships between many characteristics drawn from the past, which aids the system in increasing agricultural productivity [19]. Temperature and rainfall are two elements that affect crop output.

These time series data were subjected to Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) algorithms to improve accuracy [20]. The temperature and rainfall are predicted using historical data utilising the ARMA (Auto Regressive Moving Average), SARIMA (Seasonal Auto Regressive Integrated Moving Average), and ARMAX (ARMA with Exogenous Variables) techniques. The crop production forecast system that is based on fuzzy logic uses the model that performs the best among them. The suggested approach uses exogenous factors like cloud cover and evapotranspiration [21].

3. METHODOLOGY:

For the supplied data sets, the proposed model forecasts crop yield. By raising yields and maximising the use of available resources, integrating agriculture and ML will help the agriculture industry go further. The most important factors in predicting present performance are historical data. Several trustworthy sources, including data.gov.in, kaggle.com, and indianwaterportal.com, are used to compile historical data. The areas of Tamilnadu are where the data sets are gathered. The information includes a number of different parameters, including state, district, year, season, crop kind, area under cultivation, productivity, etc. Other datasets containing state and district details have the variable "soil type." The retrieved soil type column is combined with the primary data set. In a similar manner, the major data sets for the particular location are supplemented with temperature and average rainfall from a different dataset. The data sets

have been prepped and cleansed. The mean values are used to replace the null values. Before running the algorithms, the category characteristics are translated into labels. Categorical values in the data sets are dealt with using the one hot encoding technique. The prediction module and the fertiliser module are the two components that make up the suggested system. Numerous services are provided through mobile applications. The farmer must go through the registration process in order to register with the app. The farmer can access the mobile application services after completing the registration process. Using the attributes that have been chosen from the data sets for the particular crop, the prediction module forecasts the crop yield. Additionally, the predict module recommends the farmer who has the greatest yielding crops. The fertiliser module instructs the farmer on when to apply fertiliser. The suggested system's flowchart is shown in Figure 1.

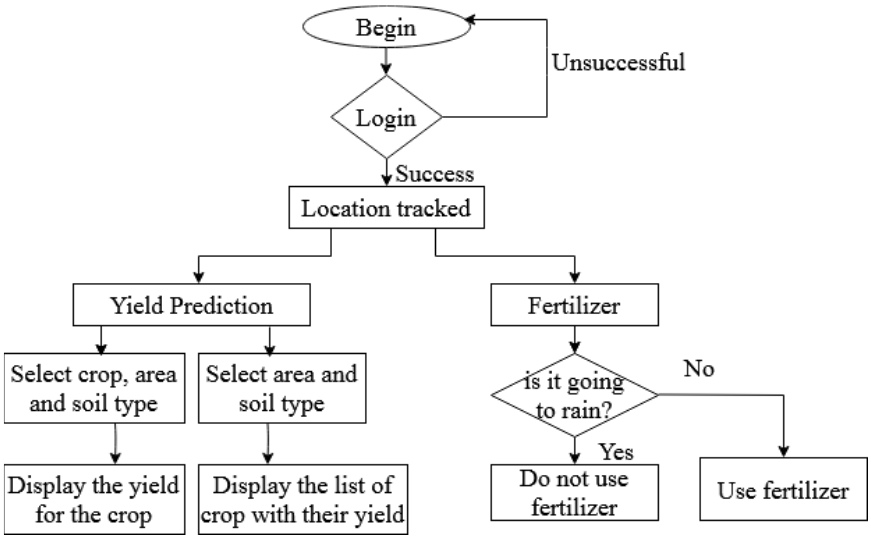

Figure 1 : Proposed system architecture

It outlines the entire procedure beginning with registration and the many services made available by the suggested system. The first service is yield prediction, either using a crop recommender system or for the chosen crop. The second function is determining when the fertiliser should be used. The user must enter the anticipated crop, the kind of soil, and the area that will be planted while using the prediction service. The technique forecasts the yield for the chosen crop. He may employ the crop recommender system this year. Only the soil type and area are required from the farmer in the crop recommender system. The system provides a list of crops together with a forecasted yield. This makes it simple for farmers to select a crop to be planted. The fertiliser application's timing is quite important. If the rain falls too soon, the farmer's time and money will be squandered. The suggested service for fertiliser usage will advise the farmer on when to apply the fertiliser. Using Open Weather API, the model forecasts rain for a specified area for the following 14 days. It advises against using fertilisers if the rainfall is greater than 1.25 mm and is considered "not safe."

4. RESULTS & DISCUSSION:

This section examines the findings from certain algorithms for the areas of Maharashtra and Karnataka. Crop type, year, season, soil type, area, and region are the factors employed by algorithms. The accuracy of the predicted crop production for each of the chosen algorithms is compared. With a 95% accuracy rate, the Random Forest method proved to be the best for the provided data set. Selected ML techniques including ANN, SVM, Multivariate Linear Regression, Random Forest, and KNN are used to forecast agricultural yield. The results of the accuracy comparison of several ML algorithms are summarised in Table 1. The results are represented graphically in Figure 2.

TABLE I : Accuracy Level of Algorithms

ALGORITHM	ACCURACY (%)
Artificial Neural Network	78
Support Vector Machine	62
Multivariate Linear Regression	76
Random Forest	92
K Nearest Neighbor	95

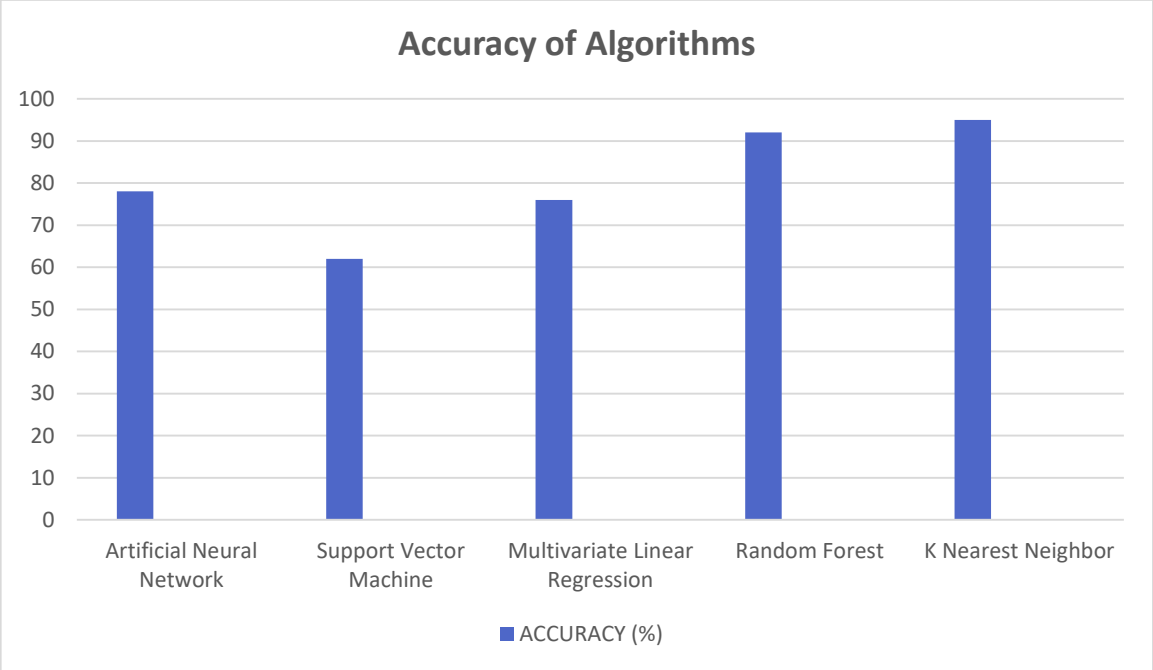

Figure 2: Accuracy level of different Algorithms

5. CONCLUSION:

The limits of current technologies and their usefulness for yield prediction were emphasised in this article. Then, a suggested system connects farmers via this system and takes them through a workable yield forecast method. Users of the system may choose from a variety of features to help them choose a crop. The built-in prediction technology aids farmers in forecasting crop yields. The built-in recommender system enables a user to investigate potential crops and their yield in order to make better informed judgements. On the provided datasets from the state of Tamil nadu many machine learning algorithms, including Random Forest, ANN, SVM, MLR, and KNN, were deployed and assessed for yield to accuracy. The accuracy of the various algorithms is contrasted. The findings show that K Nearest Neighbor, which has a 95% accuracy rate, is the best standard algorithm when applied to the supplied datasets. The suggested model looked at when fertilisers should be applied and suggested a suitable time frame. Future study will concentrate on periodically updating the datasets to provide reliable forecasts, and the procedures can be automated. The right kind of fertiliser must be provided for the crop and place in question as another functionality to be implemented.

REFERENCES:

1. Umamaheswari S, Sreeram S, Kritika N, Prasanth DJ, “BIoT: Blockchain-based IoT for Agriculture”, 11th International Conference on Advanced Computing (ICoAC), 2019 Dec 18 (pp. 324-327). IEEE.
2. Jain A. “Analysis of growth and instability in the area, production, yield, and price of rice in India”, Journal of Social Change and Development, 2018;2:46-66
3. Manjula E, Djodiltachoumy S, “A model for prediction of crop yield” International Journal of Computational Intelligence and Informatics, 2017 Mar;6(4):2349-6363.
4. Sagar BM, Cauvery NK., “Agriculture Data Analytics in Crop Yield Estimation: A Critical Review”, Indonesian Journal of Electrical Engineering and Computer Science, 2018 Dec;12(3):1087-93.
5. Wolfert S, Ge L, Verdouw C, Bogaardt MJ, “Big data in smart farming– a review. Agricultural Systems”, 2017 May 1;153:69-80.
6. Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HC, Herrero M, Howitt RE, Janssen S, Keating BA, “Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural systems”, 2017 Jul 1;155:269-88.
7. Johnson LK, Bloom JD, Dunning RD, Gunter CC, Boyette MD, Creamer NG, “Farmer harvest decisions and vegetable loss in primary production. Agricultural Systems”, 2019 Nov 1;176:102672.
8. Kumar R, Singh MP, Kumar P, Singh JP, “Crop Selection Method to maximize crop yield rate using a machine learning technique”, International conference on smart technologies and management for computing, communication, controls, energy, and materials (ICSTM), 2015 May 6 (pp. 138-145). IEEE.

9. Sriram Rakshith.K, Dr.Deepak.G, Rajesh M, Sudharshan K S, Vasanth S, Harish Kumar N, "A Survey on Crop Prediction using Machine Learning Approach", In International Journal for Research in Applied Science & Engineering Technology (IJRASET**),** April 2019, pp(3231- 3234)
10. Veenadhari S, Misra B, Singh CD, "Machine learning approach for forecasting crop yield based on climatic parameters", In 2014 International Conference on Computer Communication and Informatics, 2014 Jan 3 (pp. 1-5). IEEE.
11. Ghadge R, Kulkarni J, More P, Nene S, Priya RL, "Prediction of crop yield using machine learning", Int. Res. J. Eng. Technol. (IRJET), 2018 Feb;5.
12. Priya P, Muthaiah U, Balamurugan M, "Predicting yield of the crop using machine learning algorithm", International Journal of Engineering Sciences & Research Technology, 2018 Apr;7(1):1-7.
13. S. Pavani, Augusta Sophy Beulet P., "Heuristic Prediction of Crop Yield Using Machine Learning Technique", International Journal of Engineering and Advanced Technology (IJEAT), December 2019, pp (135-138)
14. Plewis I, "Analyzing Indian farmer suicide rates", Proceedings of the National Academy of Sciences, 2018 Jan 9;115(2): E117.
15. Nishant, Potnuru Sai, Pinapa Sai Venkat, Bollu Lakshmi Avinash, and B. Jabber. "Crop Yield Prediction based on Indian Agriculture using Machine Learning." In *2020 International Conference for Emerging Technology (INCET)*, pp. 1-4. IEEE, 2020.
16. Kale, Shivani S., and Preeti S. Patil. "A Machine Learning Approach to Predict Crop Yield and Success Rate." In 2019 IEEE Pune Section International Conference (PuneCon), pp. 1-5. IEEE, 2019.
17. Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 736-741. IEEE, 2020.
18. Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE, 2019.
19. Bang, Shivam, Rajat Bishnoi, Ankit Singh Chauhan, Akshay Kumar Dixit, and Indu Chawla. "Fuzzy logic based crop yield prediction using temperature and rainfall parameters predicted through ARMA, SARIMA, and ARMAX models." In *2019 Twelfth International Conference on Contemporary Computing (IC3)*, pp. 1-6. IEEE, 2019.