JOURNAL PUNJAB ACADEMY OF SCIENCES (Peer Reviewed Open Access Journal)
JPAS, 23:123-131 (January-December, 2023)
Available online at www.jpas.in
ISSN 2229-7014 (P)

RECENT MACHINE LEARNING ADVANCES IN AGRICULTURE

Paramveer Kaur[1], Dr.Brahmaleen Kaur Sidhu[2]

Department of Computer Science and Engineering , Punjabi University, Patiala, India[1], paramveer1067@gmail.com[1],
Department of Computer Science and Engineering., Punjabi, University,, Patiala, India [2], brahmaleen.sidhu@gmail.com[2]

ABSTRACT

The economy of India is significantly influenced by agriculture. The majority of Indians are either directly or indirectly dependent on agriculture. Therefore, it cannot be denied that agriculture plays an important role in the nation. So it is important for former to make best decision while growing any crop based on multiple factors. In this era of advancement there are number of technologies which can be used in agriculture field one of these is Machine Learning (ML). In this paper we have summarized different ML algorithms that are implemented by different researchers in their studies for Crop recommendation, yield prediction. We have concluded that Random Forest, Support Vector Machine, Naïve Bayes, Neural Network, Decision Tree, K-nearest neighbour, XGBoost(eXtreme Gradient boosting), Multivariate Linear Regression, Logistic Regression, Chi-Square Automatic Interaction Detection (CHAID) and Sliding Window non-Linear Regression comes with maximum accuracy. This prediction has been done based on various factors such as soil type, temperature, rainfall, pH value of soil, Nitrogen Phosphorus Potassium (NPK) content of soil, sowing season, porosity of soil, erosion, water holding, drainage, crop yield, crop consumed, humidity and location parameters. Furthermore, in future ML algorithms can be utilized in various sectors of Agriculture like disease detection in crops, soil suitability recommendation and prediction of retail prices

Keywords: Agriculture; Crop Recommender System; Datasets; Machine Learning; Performance Evaluation

INTRODUCTION

India has a lengthy history of agricultural development. India was recently ranked second in the world for agricultural output. In 2009, industries connected to agriculture, such as forestry and fishing, produced 16.6% of the GDP and employed nearly 50% of the workforce. Agriculture no longer makes as much money for India's GDP. Crop output is a crucial factor in determining the financial worth of agriculture. Numerous variables, such as climatic, topographical, organic, and economic concerns, have an impact on crop productivity. Farmers find it difficult to decide when and which crops to cultivate as a result of fluctuating market prices. According to a survey throughout the preceding ten years, the suicide rate in India has ranged between 1.4 and 1.8% per 100,000 people. The best crops to produce and the optimum time to plant them are topics of uncertainty for farmers. In this circumstance, crop production rates are consistently declining. Solutions have been presented to solve this issue by offering predictions on crop sustainability and suggestions based on ML models developed taking important environmental and financial factors into account.

Using ML mechanisms, a wide range of issues and difficulties can be solved in agricultural domain. Farmers can utilize ML as a tool to gather information and data, and then apply it in agriculture to increase productivity through the use of information technology. Applications for ML algorithms in agriculture include crop suggestions based on pest detection in plants, soil fertility, weed detection, crop yield cultivation, and plant disease detection based on disease identification in early

stages will recover the plant so that the crop will automatically increase. Given the importance of food quality a human health, it is crucial to reduce the use of pesticides. Intelligent agriculture must analyze how weather conditions affect the farm's environment and how long-term agricultural farming affects the soil's structure and how it results in soil erosion in addition to monitoring environmental conditions on a farm. By limiting water management, we may conserve water and use it for another crop rather than wasting it. Crop production issues can be avoided by employing ML to monitor a farm. Crop management, which includes everything from planting to harvesting, increases crop productivity.

The purpose of this study is to review the various ML Methodologies that are used to introduce an intelligent crop recommendation system by considering various environmental parameters. In addition to this a comparative study has been presented in this paper between numbers of ML algorithms which are Support Vector Machine, Random Forest , Naïve Bayes, Decision Tree, K-Nearest Neighbor, CHAID, Multivariate Linear Regression, Logistic regression, XGBoost, Neural Networks, Artificial Neural Networks with their accuracy and efficiency.

There are four sections in this paper. The first section contains introduction, in second section literature survey has been done and third section is conclusion. The references have been listed in the fourth section.

LITERATURE REVIEW

S.Pudumalar et al. in their study (A, et al. 2021), discuss about precision agriculture. Beginning with the fundamentals of precision farming, the author works towards creating a model that would enable it. The needs and strategy necessary for creating a software model for precision farming are covered in the paper. It is a cutting-edge farming method that makes use of research data on soil types, soil properties, and crop yield data gathering to advise farmers on the best crop to grow based on site-specific factors. This increases production and decreases crop selection errors. This topic is addressed in this research by offering a recommendation system using an ensemble model with majority voting approach employing Random tree, Naïve Bayes, CHAID and K-Nearest Neighbor as learners to efficiently and accurately propose a crop for the site-specific parameters. The article included a suggestion for future work that includes yield prediction and is aimed at an enhanced dataset with several features.

The purpose of the work reported in (Doshi , et al. 2018) is to offer an attempt to forecast the crop yield and price that a farmer may achieve from his property by analyzing trends in the data that have been collected in the past. Sliding window non-linear regression technique is used in order to make predictions based on the various factors that affect agricultural production such as rainfall, temperature, market prices, area of land, and previous yield of a crop. The research is carried out for a number of districts located within the Indian state of Tamilnadu. In order to help farmers' better respond to the ongoing socioeconomic crisis that many farmers are currently experiencing, the system is designed to make recommendations regarding which crops would be most suitable for them to grow. The study ends with the proposal that is integration of this with other departments working for the agricultural advancement of the nation is another way in which the system may be improved further.

Zeel Doshi et al. in the study (Gosai, et al. 2021) has taken four crops rice, cotton, sugarcane, and wheat into an account for the development of a crop recommendation system that takes into consideration the soil dataset. The soil dataset is first subjected to preprocessing, and then the ensembling method is put into action to play a significant role in the classification of the four crops. Random Forest, Naive Bayes, and Linear Support Vector Machine are the individual base learners that are used in the ensemble model. The combination method that uses the Majority Voting Technique has been used because it provides the highest level of accuracy. The efficiency gained with ensembling approach is 99.91%. As a result, the job that is being proposed

lends a helpful hand to farmers in the process of ssmaking an informed choice about the crop that will be cultivated. It results in a significant increase in crop productivity which, in turn stimulates economic growth throughout the nation.

The author in the research (Kanaga Suba Raja , et al. 2017) uses the Big Data Analytics and ML together to introduce an AgroConsultant an intelligent system that helps the Indian Farmers to choose the best crop based on the sowing season, the location of their farm, the properties of the soil, and other environmental aspects like temperature and rainfall. In addition researchers have also developed a secondary method known as Rainfall Predictor, which forecasts the amount of precipitation that will fall over the course of the next year. Decision Tree, K-NN, Random Forest and Neural Networks are used to develop the proposed system. Future modifications to the model suggested in this article might include a component that forecasts crop rotations. The selection of which crop to grow will now also rely on which crop was harvested in the previous cycle, ensuring maximum output. Finally, the study concludes that in addition to crop demand and supply, additional economic variables such as farm harvest prices and retail prices may be taken into account as components of the Crop Suitability Predictor model. This would allow for a more complete forecast that takes into account economic considerations alongside environmental and geographical ones.

Shilpa Mangesh Pande et al. in the research (Kulkarni, et al. 2018) presents a practical and approachable yield forecast method to farmers. 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 lucrative crop list may be picked using ML algorithms, and they can also forecast crop yields for user-selected crops. A few ML methods, including Support Vector Machine (SVM), Random Forest (RF), Multivariate Linear Regression (MLR), Artificial Neural Network (ANN) and K-Nearest Neighbor (KNN), are used to forecast agricultural 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 output. Furthermore, in order to make precise forecasts, future work will concentrate on periodically updating the datasets, and the procedures can be automated. The provision of the appropriate fertilizer for a specific crop and area is another service that has to be accomplished.

The study (Pande, et al. 2021) recommends a method to help farmers in crop selection by taking into account variables such as planting time, soil attributes like soil type, pH value, and nutrient concentration, climatic factors like rainfall, temperature and geographical position in terms of the state. The suggested system also includes a mechanism for yield prediction. Overall the system consist of three parts Financial Analysis, crop recommendation and sustainability prediction of crops. Linear regression and neural networks are used to implement the proposed system whereas in terms of accuracy, the outcomes are analyzed and compared with crop recommendation using K Nearest Neighbor, K Nearest Neighbor with cross validation, Decision Tree, Naive Bayes, and Support Vector Machine. In the conclusion web interface and mobile app are planned for the future so that the crop cultivation advice provided to farmers may be accessible by millions of farmers across the country.

Another work (Ray, Das and Chakravarty 2022) predicts the appropriate crop based on input characteristics such as nitrogen (N), phosphorus (P), potassium (K), soil PH, humidity, temperature, and rainfall. The research has conducted using ML methods such as Decision Trees, Nave Bayes, Support Vector Machine, Logistic Regression, Random Forest, and XGBoost are used to implement the proposed system and the results showed that XGBoost has the highest accuracy. The author concludes the study with the primary objective of the future effort is to provide a more enhanced dataset that has a greater number of attributes and to construct a model that can discriminate between healthy and unhealthy crop leaves and also, if the

crop has any illness at all, forecast which disease it is.

In research work (S.Pudumalar, et al. 2016) statistical methods such as distribution analysis, majority voting, correlation analysis, and ensembling, are applied to make 22 specific crop recommendations. To solve the crop suggestion issue, the authors offer a three-level system. The system is comprised of a data preprocessing module, a classification module, and a performance evaluation module. Correlation and density plots are used for feature analysis, and ensembling methods are then used for classification. Finally, a majority vote is taken to assess the employee's performance. To that end, the article employs ensembling methods with base learners (decision trees, random forests, Naive Bayes, and support vector machines) through the use of majority voting. In addition, final performance measures are decided by majority vote. Correlation and performance plots, as well as density and confusion matrices, are presented in a graphical format for easy interpretation. Implementing the Naive Bayes classifier resulted in an accuracy of 99.54 percent which is highest. For future research the author suggest real-time applications can be made with these classifiers in the backend to predict the best crop based on the parameters given by the user. Also, cloud-based implementation using the latest techniques of server less computing in combination with location. Using geospatial analysis to make location-based predictions can be a future research direction. Also, crop yield prediction, crop disease detection, and macronutrient recommendation can be added to the crop recommendation module to make it more accurate.

The paper reviewed in the study are summarized in the following table1.

Table 1: Summary of research studies

Study	Objectives	Machine learning technologies used	Dataset used/data sources	Features	Performance Evaluation	Gaps/Future directions
SHILPA MANGESH PANDE, DR.PREM KUMAR RAMESH, ANMOL, B.R AISHWARYA, KARUA ROHILLA, KUMAR SHAURYA (Pande, et al. 2021)	Crop Recommendation System	1.Arttifical Neural Networks 2.Support Vector Machine 3.Multivariate Linear Regression	Kaggle.com Data.gov.in Indianwaterprtal.com Maharashtra, Karnataka	1.Soil type 2.Temperature 3.Rainfall	95% using Random Forest	1.Updated Dataset 2.Fertilizer Recommendation

		4.Random Forest 5.K Nearest Neighbor				
NIDHI H KULKARNI, DR.GN SRINIVASAN, DR.BM SAGAR, DR.N K CAUVERY (Kulkarni, et al. 2018)	Create an accurate crop recommender system	Ensembling Techniques 1.Random Forest 2.Linear SVM 3.Naive Bayes	Data.gov.in Crops 1.Cotton 2.Sugarcane 3.Rice 4.wheat	1.Type of Soil 2.Soil's pH value 3.NPK concentration of soil 4.Porosity of soil 5.Avearge Rainfall 6.Sowing Season 7. Surface Temperature	99.91% average accuracy using Random Forest ,Linear SVM and Naïve Bayes	Developed System specific to four crops
S.PUDUMALARE.RAMARUJAMR.HARINE RAJASHREE C.KAVEYA, T.KIRUTHIKA, J.NISHA (S.Pudumalar, et al. 2016)	Crop Recommendation with high accuracy and efficiency	Ensembling Techniqe 1.Random Tree 2.CHAID 3.K-Nearest Neighbor 4.Naive Bayes	1.Milllet 2.Groundnut 3.Pulses 4.Cotton 5.Vegetables 6.Banana 7.Paddy 8.Soregum 9.Sugarcane 10.Corriander	1.Depth 2.Texture 3.Ph 4.Soil color 5.Permeability 6.Drainge 7.Water Holding 8.Erosion	88%	1.Improved Dataset 2.More no of attributes 3. Yield Prediction.

JPAS-2023)

S.KANAAGA SUBA RAJA,RISHI R,SUNDARESAN E,SRIJIT V (Kanaga Suba Raja , et al. 2017)	Crop Recommendation System for crop yield and price prediction	Sliding Window non-linear Regression	1.Sugarcane 2.Tomato 3.Onion 4.Wheat 5.Potato 6.Groundnut 7.Maize 8.Cucumber 9.	1.Crop Area 2.Nature of soil 3.Type of crop cultivated 4.Crop yield 5.Crop Consumed	Not specified	Comparison with other recommender systems
PRIYADHARSNI A, AUYUSH KUMAR, SWAPNEEL CHAKRABORTY, OMEN RAJENDRA POONIWALA (A, et al. 2021)	Crop Recommender System	1.Decision Tree 2.K Nearest Neighbor 3. Neural Network 4.Linear Regression 5.Support Vector Machine 6. K Nearest Neighbor with Cross Validation 7.Naive Bayes	Kaggle.com	1.Yield 2.Cost of cultivation 3.Average Market price 4.Order-state 5.NPK Content 6.Average ph 7.Max rainfall 8.Min rainfall 9.Max temperature 10.Min temperature 11.ph values	1.Neural Network 89.88% 2.K Nearest Neighbor 85% 3. .Support Vector Machine 78%. 4.Linear Regression Model 88.26% 5. Decision Tree 81% 6. Naive Bayes 82% 7 .K Nearest Neighbor with Cross Validation 88%	1.Web Interface 2.Mobile Application
ZEEL DOSHI,RASHI AGARWAL, SUBHASH NADKARNI,	Intelligent System Agroconsultant	1.Decision Tree 2.KNN	India Agricultural and Climate Data set	1.Type of Soil 2.Aquifer	1. Neural Network 91.00%(Implemented)	1.Crop Rotations

PROP.NEEPA SHAH (Doshi , et al. 2018)	Rainfall Predictor	3.Random Forest 4.Neural Network		Thickness 3. Ph of Soil 4.Thickness of topsoil 5.Precipitation 6.Temperature 7.Location Parameters	2. Decision Tree 90.20%. 3. Random Forest 90.43% 4. KNN 89.78%	
DHRUVI GOSAI, CHINTAL RAVAL, RIKIN NAYAK, HARDIK JAYSWAL, AXAT PATEL (Gosai, et al. 2021)	Crop Recommendation System	1.XGBoost 2.Naive Bayes 3.Random Forest 4.Logistic Regression 5.Decsion Tree 6.Support Vector Machine	Crops 1.Rice 2.Maize 3.Chickpea 4.Kidneybeans 5.Pigeonpeas 6.Mothbeans 7.Mungbeans 8.Blackgram 9.Lentil 10.Pomegranate 11.Banana 12.Mango 13.Grapes 14.WaterMelon 15.Muskmelon 16.Apple 17.Orange 18.Papaya 19.Coconut 20.Cotton	1.Nitrogen(N) 2.Phosphorous 3.Potassium(K) 3.Ph Value of Soil 4.Humidity 5.Temperature 6.Rainfall	1. XGBoost 99.31% 2 .Naive Bayes 99% 3. Random Forest 99% 4. Logistic Regression 95.22% 5. Decision Tree 90% 6. Support Vector Machine 10.68%.	1.Improved Dataset 2.Disease Prediction 3.Website and Mobile App

			21.Jute 22.Coffee			
RAKESH KUMAR RAY,SANJEEV KUMAR DAS,SUJATA CHAKRAVARTY (Ray, Das and Chakravarty 2022)	Smart Crop Recommender System	1.Decision Tree 2.Support Vector Machine 3.Naive Bayes 4.Random Forest	1.Pigeonpeas 2.Apple 3.Blackgram 4.Muskmelon 5.Pomegranate 6.Orange 7.Mungbean 8.Watermelon 9.Mango 10.Chickpea 11.Maize 12.Rice 13.Papaya 14.Grapes 15.Coffee 16.Lentil 17.Coconut 18.Cotton 19.Kidneybeans 20.Jute 21.Banana 22.Mothbean	1.Nitrogen 2.Phosphorus 3.Potassium 4. Average rainfall in millimeters 5.Humidity 6.pH of soil 7.Temperature	Naïve Bayes 99.54%	Location based prediction

CONCLUSION

In this paper studies conducted by different authors on ML in agriculture area mainly crop recommendation systems are analyzed ML techniques , Random Forest, Naïve Bayes, Decision Tree, Neural Network, Support Vector Machine, XGBoost, K-nearest neighbor, Multivariate Linear Regression, Sliding Window non-Linear Regression, CHAID and Logistic Regression are implemented to develop an accurate Crop Recommender System This system has developed using a number of different variables, including the type of soil, the temperature, the amount of rainfall, the soils' pH value, the soil's NPK concentration, the sowing season, the porosity of the soil, erosion, water holding capacity, drainage, crop yield, crop consumed, humidity, and location parameters. In addition, in the not-too-distant future, algorithms for ML may be applied in several areas of agriculture, including the diagnosis of diseases in crops, the suggestion of soil compatibility, and the forecasting of retail prices

REFERENCES

A, Priyadharshini, Aayush Kumar, Swapneel Chakraborty, and Omen Rajendra Pooniwala. "Intelligent Crop Recommendation System using Machine Learning." Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021)IEEE Xplore Part Number: CFP21K25-ART. 2021. 843-848.

Doshi , Zeel, Rashi Agrawal, Subhash Nadkarni, and Prof. Neepa Shah. "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms." 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). 2018.

Gosai, Dhruvi, Chintal Raval, Rikin Nayak, Hardik Jayswal, and Axat Patel. "Crop Recommendation System using Machine Learning." International Journal of Scientific Research in Computer Science, Engineering and Information Technology ISSN : 2456-3307 UGC Journal No : 64718, 2021: 554-569.

Kanaga Suba Raja , S., Rishi R., Sundaresan E., and V. Srijit . "Demand Based Crop Recommder System For Farmers." 2017 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development. 2017. 194-199.

Kulkarni, Nidhi H, Dr. G N Srinivasan, Dr. B M Sagar, and Dr. N K Cauvery. "Improving Crop Productivity Through A Crop Recommendation System Using Ensembling." 3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018. 2018. 114-119.

Pande, Shilpa Mangesh, Dr. Prem Kumar Ramesh, Anmol, B.R Aishwarya, Karuna Rohilla, and Kumar Shaurya. "Crop Recommender System Using Machine Learning Approach." Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021)IEEE Xplore Part Number: CFP21K25-ART. 2021. 1066-1071.

Ray, Rakesh Kumar, Sanjeev Kumar Das, and Sujata Chakravarty. "Smart Crop Recommender System- A Machine Learning Approach." 12th International Conference on Cloud Computing, Data Science and Engineering (Confluence 2022). 2022. 494-499.

S.Pudumalar, E. Ramanujam, R.Harine Rajashree, C. Kavya, T. Kiruthika, and J. Nisha. "Crop Recommendation System for Precision." 2016 IEEE Eighth International Conference on Advanced Computing (ICoAC). 2016. 32-36.