CROP PREDICTION USING MACHINE LEARNING: A SURVEY PAPER

Ashwani Kumar

Research Scholar, Magadh University, Bodhgaya

Dr. Bhola Shankar Barnwal

Assistant Professor, T. S. College, Hisua, Nawada

ABSTRACT

India's economic foundation is agriculture. By contributing a significant portion of domestic cultivation to ensure food security, it also has a significant impact on the Indian economy. Due to natural disasters and poor crop choice, today's farmers are all having difficulty cultivating their land. One of the most effective techniques to partially address the farmer's challenge is machine learning. However, there is a conundrum regarding the appropriate algorithm to use and the best approach to take. The tactics and algorithms utilised are surveyed in this study, along with their benefits and drawbacks.

Keywords—Machine Learning, prediction, Classification Technique.

INTRODUCTION

India's primary industry is agriculture. Agriculture has long been regarded as the primary source of supplies for meeting people's basic requirements. These days, it's quite challenging. Natural disasters, poor crop choices, and rainfall will all cause problems for farmers.

Enhancing crop prediction is one of the main goals for farmers in order to minimise cultivation losses. The development of machine learning in recent years has had a significant impact on every industry, including agriculture. For farmers, one of the main goals is to improve crop forecast and minimise cultivation losses. For various use cases, several algorithms and tools, including KNN, KNN with cross validation, Decision Tree, Regression, SVM, LS-SVM, ANN, BPN, and Deep Neural Network, have been employed in the past.

LITERATURE REVIEW

The proposed approaches for data mining comes from S. Veenadhari, Dr. Bharat Misra, and Dr. CD Singh. It is an easy-to-use website that uses the location's climatic data to anticipate the crop based on the user's preferences. They do that here by using a decision tree. Rainfall, temperature, cloud cover, wet day frequency, and 20 years' worth of agricultural yield data were acquired from various secondary sources are the parameters employed in this study. Using data from each property, we may rank the attributes. This ranking can aid in the evaluation of pertinent traits. They mostly concentrate on the soy, maize, paddy, and wheat crops in a certain district. They achieved 75% accuracy.

Shivnath Ghosh and Santanu Koley are the authors of this work. Organic matter, vital plant nutrients, and micronutrients are some of the factors employed here for soil property analysis. Here, a back-propagation neural network (BPN) is used to identify the relationship percentage between the attributes. BPN, which was programmed using reference crop growth characteristics, determines the optimal correlation proportion for these factors. First comes sampling, followed by Back Propagation Algorithm and Weight update in this machine learning system. With a specific number of neurons in the hidden layer, the results will be displayed using an artificial neural network.

Professor D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, and Chandan Mehta all employed multiple linear regression. The most profitable crop that can be grown in the given area is discovered by them. It creates a system by combining data from several sources.

This work was proposed by Konstantinos G. Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson ID, and Dionysis Bochtis. Here, they demonstrate a number of crop management techniques, including applications for yield prediction, disease detection, weed detection, crop quality, and species identification. The management system switches to real-time AI enabled programmes when ML to sensor data is present, so that they can help farmers' decision-making with strong advice. The system in this instance was mostly focused on crop prediction with a few subcategories. SVM (binary classifier) and ANN (pattern recognition) are used in this situation.

Arun Kumar, Naveen Kumar, and Vishal Vats have proposed this paper. They primarily focused on the crop sugarcane and conducted a comparative research to determine the accuracy of training the predicted model. In the production of agriculture, they employ descriptive analytics for it. Here, three different supervised approaches are applied. They are Least Squared Support Vector Machine, K-Nearest Neighbour, and Support Vector Machine. For checking, they employ three values: LOW, MID, and HIGH. Here, they discovered that SVM complexity is O(n3) and that of the least squares support vector machine is O(n).

This work was promoted by Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, and Priya R. Here is a document that examines the issues Maharastra's farmers face. utilised in the data mining methodology. It aids farmers in assessing soil quality and making crop predictions based on fertiliser and soil conditions. It compares two techniques, such as the Back Propagation Network (BPN) and the Kohonen Self Organising Map (Kohenon's SOM), which are used for supervised and unsupervised learning, respectively. The primary feature is that it accepts PH value as input along with location.

The idea for this study comes from Andrew Crane-Droesch. Climate changes are the key factor. Here, the focus was on the corn production for the Midwest region of the US. He uses deep neural networks in this instance to describe the yield modelling strategy. Here, the complicated nonlinear interactions in high-dimensional datasets are represented using a parametric structure. These have significant negative effects but less influence than standard classical statistical approaches when used under climate models.

This method was suggested by Mohsen Shahhosseini, Rafael A Martinez-Feria, Guiping Hu, and Sotirios V Archontoulis. Farmers can get the best crop recommendations by using pre-growing season crop production predictions, such as grain yields and nitrogen losses. It employs the LASSO Regression, Ridge Regression, Random Forests, and Extreme Gradient Boosting machine learning (ML) methods. Here, the major topic is (1) How do ML meta-models use pre-season data to predict maize production and N losses? How many data points are required to train machine learning algorithms to produce reliable predictions?(3) What types of input data variables are frequently employed for precise prediction? And (4) does the prediction accuracy of all ML meta-models improve? As the training dataset grew from 0.5 to 1.8 million data points across all ML models, the yield prediction error dropped by 10% to 40%, while the N loss prediction error lacked any discernible trend.

This work was proposed by Kajal Muley, Aditi Kharde, Ketki Mirashi, K.D. Yesugade, and Hetanshi Chudasama. For those involved in agriculture, predicting yield is crucial, and machine learning models and data from various sources can be used to do so. The cultivation of a chosen crop is what determines the yield of a farm, and choosing the right cultivation parameters is essential. Here, K-Means clustering is used. This technology will assist farmers in selecting the ideal crop in accordance with the existing climatic conditions, hence maximising yield rate.

S.R.Rajeswari, Parth Khunteta, Subham Kumar, Amrit Raj Singh, and Vaibhav Pandey have all proposed this paper. Here, the main emphasis is on crop production forecast, crop cost prediction, and the techniques utilised for both. These characteristics enable smart farming. After completing the feature extraction process, take the data or columns where the algorithms will be applied to determine their accuracy and plot a graph based on the resulting information. To determine the accuracy of a specific method, the classification procedure has been used. The Bayesian network is then utilised to create the statistical analysis of the attribute in the provided dataset. Following that, ANN is utilised to compare patterns with nonlinear effects and underline concepts.

This study was proposed by Ramesh A. Medar, Vijay S. Rajpurohit, and Anand M. Ambekar. Using Long-Term-Time-Series (LTTS), Weather-and-Soil Attributes, Normalised Vegetation Index (NDVI), and Supervised Machine Learning (SML) Algorithms, sugarcane yield forecasting in the Karnataka (India) region has been proposed in this article. They distinct the forecasting of yields into three stages: (i) soil-and-climate credits are anticipated for the duration of the SCLC; (ii) NDVI is anticipated using Support Vector Machine Regression (SVR) calculation by taking soil-and-climate credits into consideration; and (iii) sugarcane crop is anticipated using SVR by taking NDVI into consideration. Soil temperature, temperature, soil moisture, and precipitation are the variables employed in this. These procedures are repeated for various sets of data.

The idea for this system was proposed by Prof. Shrikant Kokate, Pavan Patil, and Virendra Panpatil. Here, it is discussed how to increase yields and pattern recognition by enhancing the system's output by including more qualities. They provide a quick overview of ML with just one attribute in other survey studies. Decision tree and Naive Bayes algorithms are combined. With the provided dataset and more variants, decision trees perform poorly, but naive bayes outperforms decision trees in certain situations. Naive Bayes and decision tree classifiers combined for classification perform better than using a single classifier model alone. These variables include soil type, soil PH value, humidity, temperature, wind, and rainfall.

The authors of this study are M. Kalimuthu, P. Vaishnavi, and M. Kishore. By using machine learning, they assist the farmer in planting reasonable crops. Here, a supervised learning approach called Naive Bayes is applied. Here, the seed information for the crops is gathered using factors that support the crops' effective growth, such as temperature, humidity, and moisture content. The suggested system consists of four main processes: gathering historical data, gathering current data, consolidating data, and gathering seed data. The Bayes Theorem estimates the likelihood of an event occurring given the likelihood of an event that has already occurred. The accuracy of the model using the Naive Bayes approach is 97%.

This paper was proposed by Merin Mary Saji , Kevin Tom Thomas, Varsha S, Lisha Varghese, Er. Jinu Thomas. By examining the agricultural region in light of the soil's characteristics, they will be able to address the challenges in agriculture. It assists farmers in increasing output and minimising loss by advising them on the best crop. A comparison of algorithms is presented in this publication. Naive Bayes, decision trees, KNN with cross validation SVM, and KNN are the key algorithms used here. The best algorithm for predicting crops was determined as a result. Testing will make use of the following algorithms: Decision Tree, Naive Bayes, SVM, kNN, and kNN with Cross Validation. The accuracy percentages that were attained were 85%, 88%, 81%, 82%, and 78%, respectively.

This work was proposed by Alexandre Barbosa, Naira Hovakimyana, and Nicolas F. Martin. In this case, it also makes advantage of the CNN architecture within the Deep Ensemble framework to boost productivity through redesign. It forecasts a probability distribution of outputs rather than a single number. Here, the maps of crop inputs are found using an optimisation approach based on gradients. With risk limitations, the net value will be maximised. The suggested model emphasises uncertainty quantification while also improving on the performance that was previously projected for it. The results of the optimisation algorithm exceed the projected net by up to 6.4%. Five input variables are used in this work: soil's shallow electro conductivity, elevation map, nitrogen and seed rate prescription maps.

The descriptions are displayed in the table.

Table: Comparative Study of Various Algorithms

YEAR	AUTHOR	PURPOSE	MENTIONED METHODS	INFERENCE
2014	S.Veenadari, Dr. Bharat Misra, Dr. CD Singh.	ML method for predicting crop productivity based on environmental factors.	Decision tree	Both classification and regression issues can be handled using decision trees.
2014	Shivnath Ghosh, Santanu Koley.	Back Propagation Neural Networks for ML for Soil Fertility and Plant Nutrient Management.	Back Propagation Neural Network(BPN) and Artificial Neural Network(ANN)	BPN determines the accurate correlation rate As a result, ANN is used for the results.
2017	D.S. Zingade ,Omkar Buchade ,Nilesh Mehta,Shubham Ghodekar ,Chandan Mehta.	Machine Learning-Based Crop Prediction System.	Multiple Linear Regression (MLR)	Multiple linear regression differs from simple linear regression in that the former uses more independent

				variables while the latter uses just one.
2018	Konstantinos G. Liakos , Patrizia Busato , Dimitrios Moshou , Simon Pearson ID and Dionysis Bochtis.	Institution for Bio-Economy and Agri-Technology, "Machine Learning in Agriculture."	Support Vector Machine (SVM) and Artificial Neural Network (ANN)	The binary classifier in this case uses SVM, and the pattern recognition uses ANN.
2018	Arun Kumar, Naveen Kumar, Vishal Vats.	Efficient Machine Learning Algorithms for Crop Yield Prediction.	Support Vector Machine (SVM) and Least Squared Support Vector Machine.	It demonstrates that, given the complexity, SVM is superior here.
2019	Andrew Crane Droesch.	Machine learning techniques for predicting agricultural yield and evaluating the effects of climate change on agriculture.	Deep Neural Network.	It is able to produce an output that is independent of the input and learn using its own method. It does not employ a database; instead, a sizable network is used. Retrieval is thus simple.
2019	Mohsen Shahhosseini , Rafael A Martinez-Feria , Guiping HU and Sotirios V Archontoulis.	Machine learning techniques for predicting maize yield and nitrate loss.	LASSO Regression, Extreme Gradient Boosting, Ridge Regression, random forests.	Farmers can receive the best crop recommendations if crop outputs like grain yields and nitrogen losses are predicted prior to the growing season.
2019	K.D.Yesugade , Hetanshi Chudasama , Aditi Kharde, Ketki Mirashi, Kajal Muley.	Utilising Unsupervised Machine Learning Algorithm, a crop suggesting system.	K-Means clustering	The algorithm iteratively associates each data point with a certain cluster based on the features provided.

2019	S.R.Rajeswari , Parth Khunteta, Subham Kumar,Amrit Raj Singh,Vaibhav Pandey.	Using machine learning prediction of smart farming	Bayesian network and Artificial Neural network (ANN).	ANN has been used to compare patterns with a nonlinear effect and to emphasise concepts.
2019	Ramesh Medar & Anand M.Ambekar.	Prediction of Sugarcane Crop using Supervised Machine Learning.	SVR, Lasso, Navïe-Bayes, and Decision Tree	Comparing NaiveBayes to the other three algorithms, it performs better. Prediction accuracy for temperature, soil moisture, and soil temperature are all higher than 80%, while it is lower for precipitation.
2020	Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate.	Machine learning algorithms are used in the crop prediction system.	Decision tree and Naïve Bayes.	Using a single classifier model is better to using a combination of the naive bayes and decision tree classifiers for classification.
2022	M.Kalimuthu ,P.Vaishnavi, M.Kishore.	Crop Prediction Using Machine Learning.	Naïve Bayes.	These results were accurate to a degree of 85%, 88%, 81%, 82%, and 78%, respectively. The most accurate algorithm is kNN with cross validation.
2022	Alexandre Barbosa, Naira Hovakimyan, Nicolas F. Martin	Using a deep ensemble of convolutional neural networks, risk-averse crop input optimisation	Convolutional Neural Network (CNN).	The results of the optimisation algorithm exceed the projected net by up to 6.4%.

CONCLUSION AND FUTURE WORKS

Here, various forms of crop prediction machine learning methods are discussed. Here, we employ a number of machine learning algorithms and compare their features to determine the optimal algorithm. In various circumstances, each algorithm has produced a different outcome. The Naive Bayesed approach provides additional marginal accuracy. In the future, it might even go so far as to recommend fertiliser, proper farming practises, and crops for the given input. Building a hardware unit with microprocessors, a DTH11 sensor, a soil sensor, and a cloud platform can create IoT systems.

REFERENCES

1. S.Veenadhari Dr. Bharat Misra Dr. CD Singh “MLapproach for forecasting crop yield based on climatic parameters”2014 International Conference on Computer Communication and Informatics (ICCCI -2014), Jan. 03 – 05, 2014, Coimbatore, INDIA.
2. Shivnath Ghosh,Santanu Koley, “ML for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks” IJRITCC, vol. 2, Issue 2,292-297,2014.
3. Prof. D.S. Zingade ,Omkar Buchade ,Nilesh Mehta,Shubham Ghodekar ,Chandan Mehta “Crop Prediction System using Machine Learning”International Journal of Advance Engineering and Research Development Special Issue on Recent Trends in Data Engineering Volume 4, Special Issue 5, Dec.-2017.
4. Konstantinos G. Liakos , Patrizia Busato , Dimitrios Moshou , Simon Pearson ID and Dionysis Bochtis “Machine Learning in Agriculture”Institute for Bio-Economy and Agri-Technology (IBO), CERTH Received: 27 June 2018; Accepted: 7 August 2018; Published: 14 August 2018.
5. Arun Kumar, Naveen Kumar, Vishal Vats, “Efficient Crop Yield Prediction Using Machine Learning Algorithms”, International Research Journal of Engineering and Technology (IRJET)- e-ISSN: 2395-0056, pISSN:2395-0072, Volume: 05 Issue: 06 | June2018 .
6. Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, Priya R L, “Prediction of crop yield using machine learning”, 2018 - International Research Journal of Engineering and Technology.
7. Andrew Crane Droesch, “Machine learning methods for crop yield prediction and climate change impact assessment in agriculture”, Published by IOP Publishing Ltd Volume: 05 | OCT-2019.
8. Mohsen Shahhosseini , Rafael A Martinez-Feria , Guiping HU and Sotirios V Archontoulis” Maize yield and nitrate loss prediction with machine learning algorithms.RECEIVED 29 April 2019 REVISED 29 October 2019 PUBLISHED 4 December 2019.
9. K.D.Yesugade , Hetanshi Chudasama , Aditi Kharde, Ketki Mirashi, Kajal Muley” Crop Suggesting System Using Unsupervised Machine Learning Algorithm” International Journal of Computer Sciences and Engineering Open Access Research Paper Vol.-7, Issue3, March 2019 E-ISSN: 2347-2693.
10. S.R.Rajeswari , Parth Khunteta, Subham Kumar,Amrit Raj Singh,Vaibhav Pandey “Smart Farming Prediction Using Machine Learning” - International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-7 May, 2019.
11. Ramesh Medar & Anand M. Ambekar, “Sugarcane Crop prediction Using Supervised Machine Learning" published in International Journal of ntelligent Systems and Applications Volume: 3 | August 2019.

12. Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate" Crop Prediction System using Machine Learning Algorithms"- International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 02 | Feb 2020 www.irjet.net.
13. M Kalimuthu ,P.Vaishnavi, M.Kishore "Crop Prediction using Machine Learning" Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT2020) IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1.
14. Kevin Tom Thomas , Varsha S , Merin Mary Saji, Lisha Varghese, Er. Jinu Thomas" Crop Prediction Using Machine Learning" International Journal of Future Generation Communication and Networking Vol. 13, No. 3, (2020), pp. 1896–1901.
15. Alexandre Barbosa, Naira Hovakimyan, Nicolas F. Martin" Risk_averse optimization of crop inputs using a deep ensemble of convolutional neural networks" Received 12 June 2022, Revised 26 August 2022, Accepted 7 September 2022.