CROP PREDICTION USING MACHINE LEARNING: A SURVEY PAPER

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.

Authors

Ashwani Kumar
Research Scholar
Magadh University
Bodhgaya.

Dr. Bhola Shankar Barnwal
Assistant Professor
T. S. College Hisua
Nawada.

I. INTRODUCTION

India's primary industry is agriculture. The primary source of resources for supplying people's basic needs has traditionally been thought of as agriculture. It might be really difficult now. Farmers may experience issues as a result of natural catastrophes, bad agricultural decisions, and rains.

One of the key objectives for farmers is to improve crop prediction in order to reduce cultivation losses. The latest advancements in machine learning have had a profound effect 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.

II. 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 thecrop 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. Usingdata 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 certaindistrict. 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 growthcharacteristics, determines the optimal correlation proportion for these factors. First comes sampling, followed by Back Propagation Algorithm and Weight update in this machine learning system. The findings will be displayed using an artificial neural network with a certain neural connections in the hidden layer.

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, et al. Here, they demonstrate a number of crop management techniques, including applications for yield prediction, disease detection, weed detection, crop quality, and species identification. The operation system switches to real-time AI enabled programmes when ML to sensor data is present, so that they can help farmers' decision-makingwith 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

K-Nearest Neighbour, Support Vector Machine, and Least Squared Support Vector Machine. For checking, they employ three values: LOW, MID, and HIGH. Here, they discovered that SVM complexityis 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 et al. With pre-growing season crop production forecasts, such as the yield of grains and fertiliser damages, farmers may acquire the best crop recommendations. It uses machine learning (ML) techniques such as the LASSO algorithm Regression, Ridge Regression, random forest training, and Extreme gradient boosts. (1) How do ML meta-models use pre-season data to predict maize production and N losses? (2) How much information is needed to train machine learning algorithms to make accurate prognosis? (3) What kinds of information found in the source data are usually used for accurate prediction? And (4) does the accuracy of all ML meta-models' predictions increase? The yield forecasting error decreased by ten to forty percent across all machine learning models as the training dataset extended from 0.5 to 1.8 million data points, but the N loss prediction error showed no obvious 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.

Ramesh A. Medar, Vijay S. Rajpurohit, and Anand M. Ambekar were the authors of the study's proposal. The forecasting of sugarcane yield in the Karnataka (India) region has been proposed in this article employing Long-Term-Time-Series (LTTS), Weather-and-Soil Attributes, Normalised Vegetation Index (NDVI), and Supervised Machine Learning (SML) Algorithms. They separate the yield forecasting into the following phases: (i) soil- and the environment credit will be expected over the life span of the SCLC; (ii) NDVI is anticipated using the support vector machine's Regression (SVR) calculation that takes soil- and the environment credits into consideration; and (iii) sugarcane crop is anticipated using SVR by taking NDVI into consideration. These variables include soil temperature, soil moisture, temperature, and precipitation. For diverse sets of data, these processes are repeated.

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 thesystem'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. Decision trees perform badly with the given dataset and additional variants, however naïve bayes beats decision trees in some circumstances. 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.

Making use of machine learning the authors Kalimuthu M. et al, 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 the surrounding temperature, water content, and moisture content. The suggested system consists of four main processes: gathering historical data, gathering current data, consolidating data, and gathering seed data. According to the Bayes Theorem, the likelihood given the possibility of an occurrence that has already happened, the likelihood of an event occurring. 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. . By advising farmers on the optimal crop, it helps them increase productivity and reduce loss. In this publication, a comparison of algorithms is offered. 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: The naive Bayes, SVM, Decision Tree, KNN, and KNN with Cross Validation are some examples of artificial neural networks. The achieved accuracy rates 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	Decision tree	Both classification andregression issues can be handledusing trees.
2014	Shivnath Ghosh, Santanu Koley.	Back Propagation Neural Networks for ML for the management of crop nutrients and soil fertility.	Neural Network with Back Propagation and Artificial Neural Network	BPN determines the accuratecorrelation rate As a result, ANN is used for the results.
2017	D.S.Zingade D.S, et al.	Crop Prediction System Based on Machine Learning.	A number of linear regressions	Multiple linear regression differsfrom simple linear regressionin that the formeruses more Independent variables the latter uses justone.
2018	G. Liakos Konstantinos et al.	"Machine Learning in Agriculture," Institution for Bio- Economy and Agri- Technology.	Support Vector Machine (SVM)and Artificial Neural Network (ANN)	The binary classifier in this case uses SVM, and the patternrecognition usesANN.
2018	Arun Kumar, Naveen Kumar, Vishal Vats.	Efficient Machine Learning Algorithmsfor Crop Yield Prediction.	Support Vector Machine (SVM) andLeast Squared Support Vector Machine.	It demonstratesthat, given the complexity, SVMis superior here.

2019	C. Droesch.	Machine learning techniques for predicting agricultural crop & evaluating the effects of climate change onagriculture.	DeepNeural Network.	It is able to produce an outputthat is independent ofthe input and learn using its own method. It does not employ a database; instead,a sizable networkis used. Retrievalis thus simple.
2019	Mohsen Shahhosseini et al.	Machine learning techniquesfor predicting maize yield and nitrate loss.	LASSO Regression, Extreme Gradient Boosting, Ridge Regression, random forests.	Farmers can receive the bestcrop recommendationsif season.
2019	K.D.Yesugade, Hetanshi Chudasama , et al.	Utilising Unsupervised Machine Learning Algorithm, a crop suggesting system.	K-Means clustering	Based on the given features, the algorithm repeatedly connects each data point with a specific cluster.
2019	S.R.Rajeswari,Parth	Using machine learning prediction of smart farming	Bayesian network	ANN has beenused to comparepatterns with anonlinear effectand to emphasiseconcepts.
2019	Ramesh Medar and Anand M.Ambekar	Sugarcane Crop Prediction Using Supervised Machine Learning.	SVR, Las so,Navïe-Bayes, and Dec isionTree	Comparing NaiveBayes While the accuracy of soil moisture, temperature, and soil temperature predictions are all higher than 80 per cent, that of precipitation is lower.

2020	Patil P, et al.	Algorithms of machine learning are used in the crop predictionsystem.	Decision tree and N aïveBayes.	Using a single classifier model is better to using a combination of the naive bayes and decision tree classifiers for classification.
2022	Kalimuthu M. et al.	Prediction of The crops Using Machine Learning.	Naïve Bayes.	These resultswere accurate to adegree of 85%, 88%, 81%, 82%, and 78%, respectively. Themost accuratealgorithm is kNN with cross validation.
2022	Alexandre Barbosa,	Using a deepensemble optimisation	Convolutional Neural Network (CNN).	The results of the optimisation algorithm exceedthe projected netby up to 6.4%.

III. 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 years to come, it could even recommend fertiliser, appropriate agricultural techniques, and crops for the 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 yieldbased 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 ofcrop 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.

