IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

CROP YIELD PREDICTION USING MACHINE LEARNING

Mr.N.Prasad[1] , B.Phani sree[2] , K.J.V.M.N.Supriya [3] , J.Satya Gowtham [4] , A.V.nagalakshmi[5]
[1]Assistant Professor, Department of Information Technology,
Sir C R Reddy College of Engineering, Eluru.
[2,3,4,5] Final Year Students, Department of Information Technology,
Sir C R Reddy College of Engineering in Eluru, Andhra Pradesh, India.

Abstract
Agriculture is very essential in a farmer's existence. Machine learning has the potential to revolutionize agriculture by altering the revenue situation by producing the optimal produce. This paper concentrates on forecasting crop yield by using different machine learning methods. Machine learning algorithm predictions will aiding producers in choosing which crop to grow to maximize output by taking into accountlike the climate, humidity, location, and so forth. This article discusses forecast of crop output using machine learning Approach and goes into depth about different approaches for disease detection and Classification.

Keywords: classification, prediction, machine learning, random forest, support vector machine.

Introduction
India's wealth is built on its agricultural sector because it is vital to both human and animal life. The expected increase in population from the present estimate of 7.9 billion to 9.8 billion by 2050 will result in a significant increase in the demand for agricultural goods which cause a notable rise in the demand for farm products. Agriculture goods will be in higher demand as the world's populace grows. necessitating efficient farmland development and an increase in crop output. Meanwhile, harvests were frequently ruined by inappropriate weather conditions as a result of global warming. Agricultural failure can be caused by calamities like drought, flooding, insect infestation, or pest attacks. Due to pests and diseases fewer products are produced, which reduces food output and contributes to food insecurity.

Prior crop forecast and yield prediction were done based on producers' expertise in a specific area. They will favor the previous or neighboring or more trend product in the nearby area only for their land, and Regarding the nitrogen, phosphate, and potassium content of the earth, they are ignorant.Given the present scenario, Reduced production, dirt toxicity, and damage to uppermost stratum are the results of crop succession and the application of inadequate fertilizers to the soil. We created the machine learning system for the cultivator's advantage while taking all of these concerns into account. The agriculture business has undergone a paradigm shift thanks to machine learning (ML).

Machine learning is not a magic trick or sorcery in the agricultural industry; rather, it is a practical tool.It pertains to a collection of very specific versions that aggregate particular factsand use specific methods to accomplish intended results.For each plot of land, the algorithm will

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

recommend the finest product. Depending on the composition of the substrate and external variables like pH, humidity, temperature, and dampness. With an objective of strengthening the farmer's scenario while taking all of the aforementioned variables to heart, we generated the system using machine learning.

To improve the findings' precision and recognition rate, modern techniques like machine learning methods have been used. Random forests are a learning technique for regression, sorting, and other tasks which requires constructing a forest of decision trees throughout training. As opposed to decision trees, random forest is able to manage both numerical and subjective data and resolves the obstacle of over-fitting of its initial information gathering.

Enhancing food production is the main objective of crop yield estimation, and numerous tried-and-true techniques are employed to achieve this.Due to its success in a number of industries, including forecasting, defect detection, pattern recognition, and other areas, ML is now utilized all over the world. The ML algorithms can additionally assist enhance the rate at which farms yield goods when there is a deficit in adverse conditions. Regardless of distracting circumstances, the crop selection method uses ML algorithms to decrease crop output production losses.

Literature Survey

To increase farmer profit and agricultural quality, Ashwani Kumar Kushwaha outlined harvestingTechniques for foresight and suggests appropriate crops. In this paper, they use the Hadoop framework and a farming programme to collect enormous amounts of data (soil and meteorological data) for predicting agricultural production. Repository data will predict agricultural suitability for particular circumstances as a result, improving harvest superiority.

Girish L anticipates farming operations and rainfall utilizing machine learning. The efficacy of distinct algorithms for machine learning, notably exponential regression, SVM, The kernel neural network tactics and others, coupled with different strategies for machine learning to gauge rainfall and harvesting, will be explored in this piece of research.

Rahul Katarya articulates the numerous methods for machine learning relied upon enhancing farming operations. In this composition,An assortment of artificial intelligence avenues including machine learning algorithms and large-scale data analysis for precision gardening.They go over KNN-based, ensemble-based, neural network, and other types of recommender systems used in agriculture.

Ananthara, M. G. et al. (2013, February) suggested the CRY algorithm for crop output forecast using beehive clustering methods for agricultural datasets. They considered factors which might involve merchandise variety, kind of dirt, and acidity level of the soil, temperature, and berry sensitivity. They mainly studied the output of maize,legumes, and cultivation in country. When their demonstrated process was put up against the C&R tree algorithm, it defeated it by 90%.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Farm yield forecast clustering kernel technique was the centre of a novel, intelligent framework developed by Awan, A. M., and colleagues in April 2006. They considered variables such as woodlands elevation, ambient temperatures, and moisture in that geographical region. They employed the weighted k-means kernel approach with aerialparameters regarding groves of oil palms.

Proposed System

The proposed system uses the Naive Bayes classifier, a supervised learning technique with four levels, to calculate and estimate crop production for the suitable environment in phenomena such

A. Data Collection

For data collections, information is gathered from various sources and modified. And comprehensive evaluation is done using the data. The data is available for up to 10 years through with a variety of abstract online sources, including Kaggle, Google Weather, and data government. Inseries. For harvest projections and increased agricultural yields, data types such soil characteristics, meteorological conditions, and seed information are used.

B. Preprocessing Step

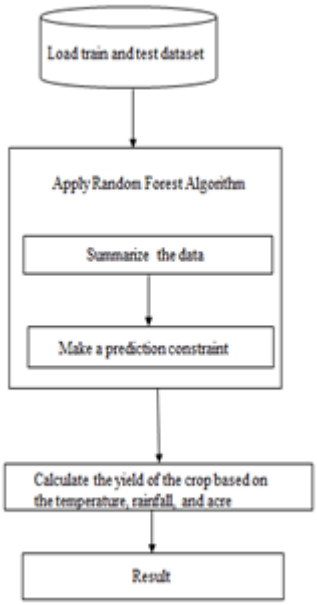
Data pre-processing the information is regarded as a crucial stage in the machine learning process. Preprocessing include adding the right collection of data, the null values, and capability extraction. The structure of the data set matters to the analysis process. Python programming will be used to import the data gathered in this stage into the Google Colab platform in order to get the required results.

C. Feature Extraction

The quantity of information required that must describe a large collection of information would be less if the characteristics were extracted. The completing the training data gathering is determined by the soil, crop, and weather characteristics gathered throughout the preparation phase. This method chooses the features according to the correlation vector, choosing the characteristics with the highest correlation value as a key predictor of yield.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022



D. Data Prediction

The data must first be divided into train dataset and test dataset before moving on to this stage. The Nave Bayes Gaussian classifier is used to develop the information using the inputs as well as the outputs that are currently accessible. The data are examined in the testing stage to see whether the model's accuracy is satisfactory. Then, a machine learning method forecasts the new data.

Table I. Accuracy of Models Applied

Model	Accuracy(in percentage)
Random forest	92.03%
Logistic regression	89.62%
Support vector machine	75.82%

It can be noticed from Table I .The random forest classifier outperformed all the other approaches in terms of accuracy.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

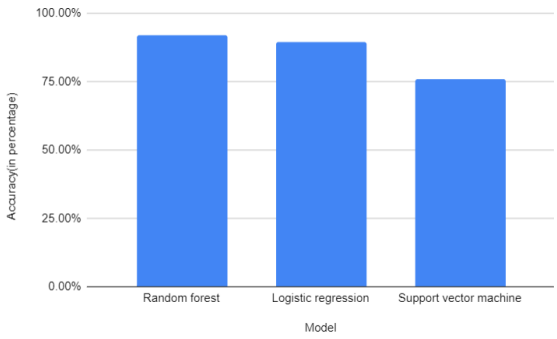


Fig 2. Several algorithms' Mean Absolute Error

Fig 2. Demonstrates that the Random Forest Regressor has the lowest mean absolute error of any machine learning method. Hence, the best method for predicting crop yield

Conclusion

As our farmers are now not making efficient use of technological innovation and assessment, there's a chance that they may choose the wrong crops to cultivate, which would reduce their revenue. We originated a fatal agronomist-friendly instrument with a an animated user interface that will forecast what harvest is an optimal fit and greatest matchto acquire particular piece of land in order to lessen these kinds of losses. Also, this system will include details on the nutrients that must beadded, the seeds that must be used in growing, the anticipated yield, and the selling price. As a consequence, producers are more inclinedthroughchoosewise crops, and new concepts will help the sector.

Future Scope

By supplying GPS coordinates for a field, we must gather all necessary data. By acquiring access to the federal pouring predicting appliance moreover we can anticipate extracts simply presenting our coordinates to GPS.Inevitably, a strategy ought to be developed. to avoid food shortages and surpluses. In the future, we will be able to identify the most efficient algorithm based on their accuracy measures, which will aid in the selection of an efficient algorithm for agricultural output prediction. An illness prediction system can be built using a deep learning architecture. Integrating image processing and deep learning methods demonstrated greater promise in illness forecast systems. More research into these methods is needed to achieve a superior prediction system.

References

[1]Prof D.S Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, and Chandan Mehta, "Crop Prediction System Using Machine Learning."

[2] Swetabhattachrya, Ashwani kumar Kushwaha, "crop yield prediction using agro algorithm in hatoop."

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

[3] Girish L, Gangadhar S, Bharath T R, Balaji K S, and Abhishek K T "Machine Learning for Crop Yield and Rainfall Prediction in Tumakuru District."
[4] Rahul Katarya, Ashutosh Raturi, Abhinav Mehndiratta, Abhinav Thapper "Impact of Machine Learning Techniques in Precision Agriculture".
[5] "Impact of Machine Learning Techniques in Precision Agriculture," Rahul Katarya, Ashutosh Raturi, Abhinav Mehndiratta, and Abhinav Thapper.
[6] "Prediction of Rainfall Using Support Vector Machine and Relevance Vector Machine," Pijush Samui, Venkata Ravibabu Mandla, Arun Krishna, and Tarun Teja.
[7] Himani Sharma and Sunil Kumar, "A Survey on Decision Tree Classification Algorithms in Data Mining."
[8] N. Gandhi, L. J. Armstrong, O. Petkar, and A. K. Tripathy. (2016, July).
Support vector machines are being used to forecast rice harvest yields in India. The 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) was held in 2016. (pp. 1-5). IEEE.
[9] N. Gandhi, L. J. Armstrong, and O. Petkar. (2016, July). Prediction Rice crop yield prediction in India using support vector machines. In 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1-5). IEEE.
[10] Gandhi, N., Armstrong, L. J., &Petkar, O. (2016, July). Proposed decision support system (DSS) for predicting Indian paddy harvest output. IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR) was established in 2016. (pp. 13-18).
[11]Bang, S., Bishnoi, R., Chauhan, A. S., Dixit, A. K., & Chawla, I. (2019, August). Crop yield prediction based on fuzzy logic using temperature and rainfall factors forecasted by the ARMA, SARIMA, and ARMAX models. The Twelfth International Conference on Contemporary Computing (IC3) will be held in 2019. (pp. 1-6).
[12] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., & Chaudhari, A. N. (2018, August). Crop Yield Prediction Using Data Analytics and a Hybrid Approach. The Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) was held in 2018. (pp. 1-5).
[13] Gandge, Y. (2017, December). A examination of different data mining methods for predicting crop yield. International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT) took place in 2017. (pp. 420-423).
[14] Gandhi, N., Petkar, O., & Armstrong, L. J. (2016, July). Rice crop yield prediction using artificial neural networks. In 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR) (pp. 105-110). IEEE

Mr. N.Prasad is working as an Asst. Professor, in IT Department, Sir C. R.Reddy College of Engg, Eluru, A.P, India. He has received his B.Tech(CSE) from SVH College of Engineering, Machilipatnam and M.Tech (CSE) from Jawaharlal Nehru Technological College of Engineering, Hyderabad. (JNTUH University), A.P, India. Pursuing PhD. at Andhra University. His research interests include Image Processing, Computer Networks, Data Mining, Network Security, Machine Learning, and Internet of Things(IoT).

ISSN PRINT 2319 1775 Online 2320 7876

Research paper© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

B.Phani sree is a student of Sir C. R.Reddy College of Engineering, Eluru, A.P, India in the Department of Information Technology Her research interests include Image Processing, Data Mining, Network security, Web security, Software Engineering and Machine Learning.

K.J.V.M.N.Supriya is a student of Sir C.R.Reddy College of Engineering, Eluru, A.P, India in the Department of Information Technology. Her research Interests include Image Processing, Data Mining, Networks Security, Web security, Software Engineering and Machine Learning.

J.Satya Gowtham is a student of Sir C. R.Reddy College of Engineering, Eluru, A.P, India in the Department of Information Technology. His research interests include Image Processing, Data Mining, Networkssecurity, Web security, Software Engineering and Machine Learning.

A.V.Naga Lakshmi is a student of Sir C. R.Reddy College of Engineering, Eluru, A.P, India in the department of Information Technology. Her research interests include Image Processing, Data Mining, Networks Security, Web security, Software Engineering and Machine Learning.