iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 7 **Issue:** IV **Month of publication:** April 2019

**DOI:** https://doi.org/10.22214/ijraset.2019.4542

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887*
*Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com*

# A Survey on Crop Prediction using Machine Learning Approach

Sriram Rakshith.K[1], Dr. Deepak.G[2], Rajesh M[1], Sudharshan K S[1], Vasanth S[1], Harish Kumar N[2]
[1, 3, 4, 5]*Student,* [2, 6]*Assistant Professor, Department of CSE, DSCE, Bangalore*

***Abstract:*** ***This paper is mainly focused on the techniques and measures taken to improve farming by inculcating the technical knowledge and developments in order to make the agricultural sector more reliable and easy for the farmers by predicting the suitable crop by using Machine learning techniques by sensing parameters like-- soil, weather and market trends. Parameters considered are PH, Nitrogen-phosphate-potassium contents of soil, temperature, rainfall and humidity. We consider Artificial Neural Network, Information Fuzzy Network and other Data Mining Techniques. The complete research comes up to a conclusion that Artificial Neural Network is the suitable technique for our project.***
***Keywords: Data mining techniques, crop prediction, weather prediction***

## I. INTRODUCTION

Agriculture plays an important role in every individual's life. Being the backbone of India, Agricultural sector has been improving by the needs of the public, as the technology improves. These developments are very much necessary to meet the needs of every individual with the rapid growth of population and considering various situation of climatological phenomenon affecting local weather conditions which has a direct impact on the crop yield, researchers have explored the relationship between the climatological phenomenon and crop yield. For better crop yield the Artificial Neural Network have demonstrated to be an effective tool for modeling and prediction. These technological aspects pressures up to improve the use of technology in the field of agriculture [1] to find out the better crop yield.To get information or data on agricultural field such as soil condition, weather condition, plant physiology and several process take place in field can achieved by using different sensors, satellites etc. These datasets are extremely helpful when it comes to the agricultural production. Machine Learning algorithms can be used to perform these tasks and help to predict the type of the crop that can be planted [2], [3], [4].

In Agricultural practices all the work is done in field and to forecast the yield the data will be analyzed by machine learning with remote sensing data over the farm. NVDI ( Normalized Difference Vegetation Index) is a technique which retrieves vegetation activity using near-infrared and red spectral channels. Agricultural production is affected by harmful environmental factors. A seasonal variation in the crop is observed due to the environmental impacts.we need use the advanced agricultural resource management techniques to improve the crop yield.Due to the increasing amount of data that is being collected, we use machine learning to improve our crop prediction.

## II. TECHNIQUES USED IN PREDICTIONS

1) *Artificial Neural Network:* Artificial Neural Networks, as the name suggests "neural" is brain-inspired word. It works as the same way as the human brain works. In Neural networks it consist of input,output and hidden layers layers,where the neurons are the input given to the ANN and it is performed by hidden layers by some units and it is used by output layers to produce output. The accuracy of the neural networks increases as the data increases. They adapt to the complexity without knowing the principles of underlying layers. In artificial intelligence and machine learning algorithms like ID3 and other optimizing algorithms are used in tomato crop detection [6]. Tomato is widely used crop around the globe, its cultivated in almost all parts of the world. To design a expert system for Tomato crop the taken help from the computer engineers to design and program and Agriculture scientist and the expert who have more knowledge in tomato cultivation.In maize cultivation, the machine learning techniques are used [7]. Corn is also a popular crop and a main source of cereals along with rice genotypes which is adapted and well suited in drought situation which has to be grown under controlled situations and marginal law has to be implemented.
2) *Information Fuzzy Network:* Crop prediction and analysis is done through neural network [8]. The inputs are Soil moisture content, ground biomass and repository organ for Neuro-fuzzy Inference system. There are other problems like forecasting yield in a remote sensing area and goes long behind in time. The design of the algorithm is in such a way that it leaves behind a year and uses the rest of the data. The deviation is determined by comparing the yield with the one that is left out.[9] The study conducted to consolidate the aspect of incomplete information, used the web based decision support system where the fuzzy logic advancement of an agricultural used because the effect of climate development is considered to be a major part in the

field of agriculture and crop prediction in Malaysia. Therefore, the synthesis o data and heterogeneity of data is important because they influence the decisions for incomplete knowledge to bring out more transparency and acceptable information for easier user communication. The detailed information about soil , rainfall arrangement will be taken into consideration to bring about accomplished results.

3) *Data Mining Techniques:* This technique is used to analyze the collected data to give insights to make decision.By using the information or result obtained by the data mining techniques we can reduce the risk associated with the agriculture by predicting the crop yield more precisely to harvest.To provide efficient result the Data mining technique required large amount of data to analyze such a way in yield prediction it requires more data that related crop yield.Weather data.soil properties and agricultural statistics etc[10].

Data mining techniques are divided into two groups:

*A.* Classification
*B.* Clustering.

Classification techniques is uses the samples classified datasets to provide the categorize information. This datasets are used to train the classification technique.If there is no training set, then it is difficult to categorize the data. In this type of cases,the unknown datasets are used to split into clusters and it is known as cluster method.
clustering method can be used to split a set of unknown data into clusters.[11].
There are seven different methods in data mining procedures [12]:

1) *Data Cleansing:* Data Cleaning is very domain specific. The data is inconsistent, always varying and quite trivial nowadays. Dirty or unwanted data in data warehouse should be cleaned to keep high data quality [25].
2) *Data Integration:* Data Integration is to consolidate the inconsistent data to consistent data. It should resolve the problems such as naming conflicts and inconsistancies [26].
3) *Data Selection:* Data Selection deals with the choosing the right data for the process.
4) *Data Transformation:* Data Transformations are transformation of datasets mathematically [27]
5) *Data Mining:* Data Mining can be classified into two techniques - Predictive and Description [28]. Either of the two methods are used in Data Mining.
6) *Pattern Estimation:* Recognizing the pattern according tot he needs of the data mining and its classifications is important.
7) *Knowledge Display:* Display of knowledge

## III. LITERATURE SURVEY

Sujatha et al describes How the old agricultural information can be used to describe the future prediction of crops and yeild. It also suggests the farmers about what type of crop can be grown using the weather stattion information and provides the suitable information to prefer the accurate season for excellence farming. Data mining techniques are discussed in detail. [16].
Kushwaha et al describes the prediction of crop using IoT with the suitable climatic conditions and the possibilities of improvement and its application. They have used the Hadoop file system[17].To build a prediction system for crops and to detect the pests the classification,analysis and prediction algorithm is used.
Fathima et al describe the different mining techniques to study crops that are quantitative and correlation them for interseason growth.Clustering large data is a challenge, so k means algorithm is used to manage large data. Appropriate algorithm is used to determine the crops are selected as frequent item set.And they focus on the government policies and the cropping practices of frames.[18]
Veenadhari et al describes important role that performed by data mining methods in agricultural field.They have presented the different ML algorithms such as k means,SVM,ANN etc. The crops were predicted mainly based on climatic features which gives accuracy score of about 95% with the C4.5 algorithm [19].
Sellam et al describes the various factors that involved in environmental parameters which influence the yield of the crop are Area Under Cultivation, Annual Rainfall, and Food Price Index and establish the relationship among these parameters. The infliction on the crop yield is analyzed by using various environmental factors and Regression Analysis (RA), Linear Regression (LR) Algorithm.[20]
Raorane et al describes by using the different data mining techniques how to improve the crop production.And the techniques they have used for classification such as ANN,SVM and k means etc.[21]

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887*
*Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com*

Kaur et al uses the different data mining technique achieve the high accuracy in price prediction.The data model is build to predict the price.They have collected the market price of the tomato and build the model and predict the price by using BP Neural Network and it is simulated using MATLAB.[22]

Ankalaki et al studied the clustering algorithm such as DBSCAN nnd AGNES.And to predict the crop yield Multiple Linear Regression is used.[23]

Gayatri et al describe the better yield for the farmers to perform this operation they have collected the data from different technologies like IOT and web services and it can handle large amount o f data.GPS is used to capture the agriculture field images and it is stored in repositories along with its location.And it is communicated through cloud.

## IV. METHODOLOGY

Data mining has two main techniques, they are classification and clustering. The future data are classified and predicted using classification and prediction which are the two ways of analyzing data . The goal is to set the accuracy of testing Hugh than the accuracy of training of classification algorithm. The three methods of data mining are Supervised learning, Unsupervised learning and Semi-supervised learning. The various classification techniques that are used to find knowledge are described below: [13].

### *A. Naive Bayes*

Naive Bayes works on the principle of Bayes theorem which uses hypothesis of strong anatomy. This algorithm strongly depends on the probability models. It works at its fullest in a supervised learning environment . The major advantage of this classification algorithm is that , it requires minimum training data for evaluating the mean and varience of variables that are important for classification.The formula for Bayes theorm is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

### *B. J48*

J48 algorithm is an implementation of java algorithm. Decision tree analysis is mainly used in C4.5 algorithm. A set of labelled data inputs are used to build a decision tree and it can be validated . This Algorithm is is an extension of ID3 algorithm. The J48 algorithm uses a process called counting gains. It involves measure of disorder which is called as “entropy”. The Entropy y is calculated by:

$$\boldsymbol{Entropy}(\vec{y}) = -\sum_{j=1}^{n} \frac{|y_i|}{|\vec{y}|} \log\left(\frac{|y_i|}{|\vec{y}|}\right)$$

$$\boldsymbol{Entropy}(j|\vec{y}) = \frac{|y_j|}{|\vec{y}|} \log\left(\frac{|y_j|}{|\vec{y}|}\right)$$

And Gain is

$$\boldsymbol{Gain}(\vec{y}, j) = \boldsymbol{Entropy}(\vec{y} - \boldsymbol{Entropy}(j|\vec{y}))\ |$$

### *C. Random Forest*

An algorithm known as Random Forest is one of the ensemble techniques for classification and regression tasks which constructs a decision tree .It is best suitable for handling huge amount of data with perfection.

### *D. Artificial Neural Network*

Neural Networks are developed using the principle of biological neural system. Neural Networks is a framework for different machine learning algorithms to work together and process data but not an algorithm itself. The theoretical and computational neuroscience are consists of the methods that are used to design neural network by using Central Nervous System (CNS)

### *E. Decision Tree*

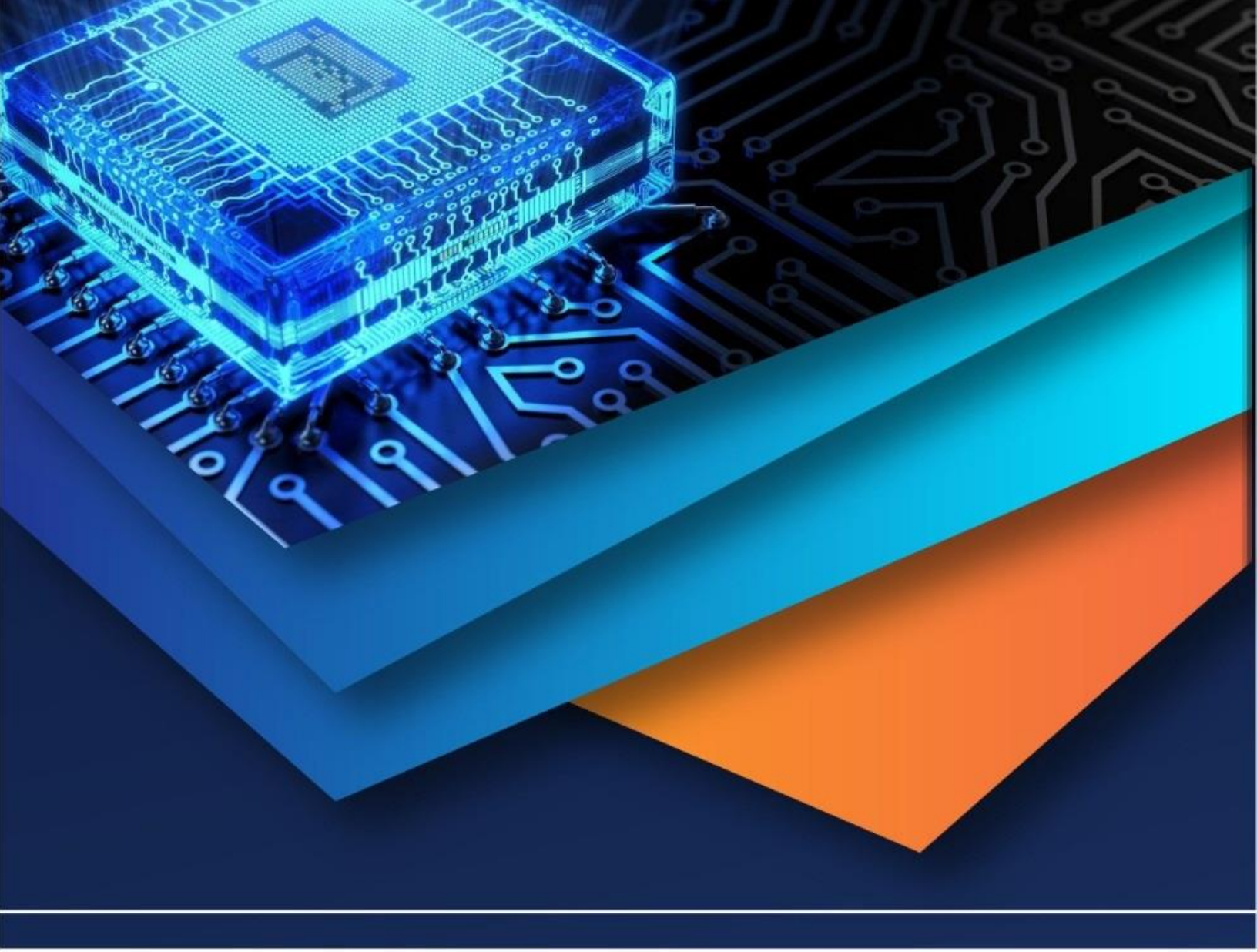
Decision Tree is one of the best data mining technique. Geospatial Decision Support System was a success by the use of mining technique which is related risk management . Decision tree is focused on learning a rule rather than the instance of observation.[14][15].

## V. CONCLUSION

We can conclude now that the task of prediction of crops can be achieved using different methodologies as discussed above. We can also say that ANN's give us a better, accurate predictions. Hence by using the soil, weather and market prices, we can build a model, having a workflow as shown above, that can provide accurate predictions about the crop yield suitable for a particular region. These are perfomed by using Artificial Neural Network (ANN).

## REFERENCES

[1] FAO, "Wfp (2015), the state of food insecurity in the world 2015. meeting the 2015 international hunger targets: taking stock of uneven progress," Food and Agriculture Organization Publications, Rome, 2016.
[2] S. S. Dahikar and S. V. Road, "Agricultural crop yield prediction using artificial neural network approach," International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, vol. 2, no. 1, pp. 683–686, 2014.
[3] B. Ji, Y. Sun, S. Yang, and J. Wan, "Artificial neural networks for rice yield prediction in mountainous regions," The Journal of Agricultural Science, vol. 145, no. 3, pp. 249–261, 2007.
[4] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. ACM, 2018, p. 50.
[5] A. Van de Griend and M. Owe, "On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces," International Journal of remote sensing, vol. 14, no. 6, pp. 1119–1131, 1993.
[6] Babu MSP, Ramana Murty NV, Narayana SVNL. A web based tomato crop expert information system based on artificial intelligence and machine learning algorithms. International Journal of Computer Science and Information Technologies. 2010; 1(3):1–5
[7] Ornella L, Cervigni G, Tapia E. Applications of Machine Learning for Maize Breeding. In: Venkateswarlu B, Shanker AK, Shanker C. Book chapter of Crop stress and its management: Perspectives and Strategies, Springer, New York, USA. 2012; 1–29.
[8] Stathakis D, Savin I, Negre T. Neuro-fuzzy modeling for crop yield prediction. The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences. 1994; 34:1–4.
[9] Salleh MNM. A Fuzzy Modelling of Decision Support System for Crop Selection. IEEE Symposium on Industrial Electronics and Applications (ISIEA2012), Bandung, Indonesia. 2012; 17–22.
[10] Maria Rossana C.de Leon,Eugene Rex L.Jalao, "A prediction model framework for crop yield prediction," Asia Pacific Industrial Engineering and Management System.
[11] Ramesh A. Medar,Vijay S. Rajpurohit, "A survey on Data Mining Techniques for crop yield prediction", International Journal of Advance Research in Computer Science and Management Studies.Volume 2 , Issue 9, Sept 2014.
[12] R. Sujatha, Dr. P.Isakki,A Study on Crop Yield Forecasting Using Classification Techniques, 978-1-4673- 8437-7/16/$31.00 ©2016 IEEE.
[13] Veenadhari, S., Bharat Misra, D Singh, "Data mining Techniques for Predicting Crop Productivity – A review article", IJCST, International Journal of Computer Science and technology, march 2011.
[14] Georg Ruß, Rudolf Kruse, Martin Schneider, and Peter Wagner. "Estimation of neural network parameters for wheat yield prediction" In Max Bramer, editor, Artificial Intelligence in Theory and Practice II, volume 276 of IFIP International Federation for Information Processing, pages 109–118. Springer, July 2008.
[15] Iv´an Mej´ıa-Guevara and ´Angel Kuri-Morales. "Evolutionary feature and parameter selection in support vector regression". In Lecture Notes in Computer Science, LNCS, volume 4827, pages 399–408. Springer, Berlin, Heidelberg, 2007.
[16] Sujatha, R., Isakki, P., "A study on crop yield forecasting using classification techniques", International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), pp.1-4, 2016.
[17] Kushwaha, A.K., SwetaBhattachrya, "Crop yield prediction using Agro Algorithm in Hadoop", International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 5- No2, pp.271-274, 2015.
[18] Fathima, G.N., Geetha, R., "Agriculture Crop Pattern Using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Engineering, Vol. 4, Issue 5, pp.781-786, 2014.
[19] Veenadhari, S., Misra, B., Singh, C.D., "Machine learning approach for forecasting crop yield based on climatic parameters", International Conference on Computer Communication and Informatics, pp.1-5, 2014.
[20] Sellam,V., Poovammal, E., "Prediction of Crop Yield using Regression Analysis", Indian Journal of Science and Technology, Vol. 9(38), pp.1-5, 2016.
[21] Raorane, A.A., Kulkarni R.V., "Data Mining: An effective tool for yield estimation in the agricultural sector", International Journal of Emerging Trends & Technology in Computer Science(IJETTCS), Vol. 1, Issue 2, pp.75-79, 2012.
[22] Kaur, M., Gulati, H., Kundra, H., "Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications", International Journal of Computer Applications, Vol. 99– No.12, pp.1-3, 2014.
[23] Ankalaki, S., Chandra, N., Majumdar, J., "Applying Data Mining Approach and Regression Model to Forecast Annual Yield of Major Crops in Different District of Karnataka", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, SpecialIssue 2, pp.25-29, 2016.
[24] Gayatri, M.K., Jayasakthi, J., Anandha Mala, G.S., "Providing Smart Agricultural Solutions to Farmers for better yielding using IoT", IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR), pp.40-43, 2015.
[25] 2009 ISECS International Colloquium on Computing, Communication, Control, and Management
[26] IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN : 2278-0661, p-ISSN : 2278-8727 PP 67-71
[27] http://periodicos.uem.br/ojs/acta ISSN on-line: 1807-8621 Doi: 10.4025/actasciagron.v40i1.35300
[28] Samiddha Mukherjee et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (5) , 2015, 4663-4666



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY