PREDICTION OF CROPS USING DEEP LEARNING – A SURVEY

[1]MD Nasreen, [2]M Arshi, [3]Vedanabhatla Anusha, [4]K. Sahithi

[1]Assistant Professor, [2]Assistant Professor, [3]Assistant Professor, [4]Assistant Professor
[1]Department of computer science,
[1]Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

Abstract: **Agriculture plays a major part in economic growth in the region. Climate and other changes in the climate have become a significant factor in agriculture Machine learning (ML) is an important approach to seeking realistic and efficient solutions for this problem. Crop yield Prediction includes forecasting crop yield from the past data like weather parameters, soil parameters and historical crop yield. This article focuses primarily on the techniques and steps taken to boost farming by inculcating technological expertise and innovations to allow the farming sector most efficient and simpler for farmers by forecasting the correct crop using machine learning approaches by sensing parameters such as soil, environment and market patterns. The parameters assumed are soil, temperature, rainfall and humidity PH, Nitrogen-phosphate-potassium content.**

Keywords: **Crop Yield, Machine learning, Forecasting, Prediction**

I. INTRODUCTION:

Agriculture forms the foundation of every economy. In a country including India, which is experiencing ever-growing demand for food due to the increasing population, growth in the agricultural field are required to satisfy the requirements. Agriculture is considered an important and foremost culture practiced in India from ancient times. The ancients plant the crops in their own land and have been adapted to their purposes. Thus, other people like humans, animals and also birds have cultivated the crops on which they have utilised them. The greenish goods created in the land that the creature took lead to a safe and well-being life. The farming field is increasingly declining, after the advent of new revolutionary technologies and strategies. As a result, plentiful innovations are focused on the production of artificial products which are synthetic products where life is unsafe.

PROBLEM DEFINITION

Modern people today have no knowledge of growing the crops in the exact period and at the right place. Those cultivating techniques also change the seasonal conditions over the major assets likewise soil, water and air that lead to food insecurity. And counting all the issues and conditions including environment, temperature, and different elements, there is no reasonable solution and technology to sort out the issue we are facing. In our country, there are many paths to improve agricultural economy. There are various ways to grow and boost crop yields and crop quality. Data mining is also useful to estimate crop yields. In normal, data mining is the method of analysing and summing up data from various viewpoints into usable knowledge.

TECHNIQUES USED IN PREDICTION

1) ARTIFICIAL NEURAL NETWORK:

Artificial Neural Networks is a term influenced by the brain, as the name suggests "neural." It functions as the workings of the human brain. In Neural networks it consists of layers of input, output and hidden layers, in which the neurons are the input supplied to the ANN and maybe some units use hidden layers to generate output. Neural networks are getting more reliable as the data increases.

We respond to the challenge, without understanding the underlying layers values. Tomato crop detection [6] uses in artificial intelligence and machine learning algorithms such as ID3 and other optimizing algorithms. Tomato is commonly used crop across the globe, grown in nearly every corner of the world. The machine learning techniques are used in maize cultivation to design an expert framework for tomato cultivation by the computer engineers to design and software and the agricultural scientist and the expert who have more expertise in tomato cultivation. [7]. Corn is also a common crop and a major source of cereals together with rice genotypes that are adapted and well adapted in a situation of drought that has to be cultivated under managed situations and limited legislation has to be enforced.

2) INFORMATION FUZZY NETWORK:

The prediction and analysis of crops is achieved by neural networks [8]. The sources for Neuro-fuzzy Inference framework are such as soil moisturising content, field biomass, and repository organ. There are many other issues, such as forecasting yield in a remote sensing field, and it goes well behind time. The algorithm's design here is that it left behind a year, using the rest of the data. The divergence is calculated by the comparison of the yield to the one left out[9].The research conducted to integrate the dimension of imperfect knowledge, used the internet-based decision support system in which the fuzzy logic advancement of an agricultural use is known to be a significant part of India’s farming and crop prediction impact. Moreover, data synthesis and data heterogeneity are relevant as they influence decisions about incomplete knowledge in order to generate further clarity and appropriate data for better interaction with users.

3) DATA MINING TECHNIQUES:

This methodology is often used to analyze the data collected and provide inputs for decision taking. By using the knowledge or findings obtained through the data mining approaches, we can decrease the risk associated with agriculture by predicting the crop yield and harvest more accurately

In order to produce an effective result, the data mining technique needed massive quantities of data in order to evaluate a way in yield prediction it needs certain information relevant to crop yield.[10].

In the present, data mining techniques can be categorized into:

1. Clustering
2. Classification

Classification methods are used to provide categorization information using the samples classification data sets. Such datasets are used to train the method of classification. When there is no training set then categorizing the data is difficult. For these instances, the unknown datasets are used to break into clusters and are regarded as cluster form.

clustering method can be used to segment a group 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.

2) Data Integration: Data Integration is to consolidate the inconsistent data to consistent data. It should resolve the problems such as naming conflicts and inconsistencies.

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

5) Data Mining: It can be divided into two techniques - Predictive and Description Either of the two methods are used.

6) Pattern Estimation: Recognizing the pattern according to the needs of the data mining and its classifications is important.

7) Knowledge Display: Display of knowledge.

II. LITERATURE SURVEY

Sujatha et al describes How the old agricultural information can be used to describe the future prediction of crops and yield. It also suggests the farmers about what type of crop can be grown using the weather station 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 identify the various mining methods used to examine quantitative crops and correlate them for inter seasonal development. Big data clustering is a task, so k means methodology is being used to handle huge data. Suitable algorithms are being used to decide the crops are chosen as regular items set. Then they concentrate on government policy then frame crop activities [18]

Veenadhari et al define the essential role played in the farming sector by data mining techniques. They introduced various ML algorithms like k means, SVM, ANN, etc. The crops were predicted based primarily on atmospheric characteristics, but with the C4.5 algorithm gives an accuracy score of around 95 percent [19].

Sellam et al Sellam et al explains and defines the relationship between these criteria and the various factors involving in atmospheric conditions which affect the crop yield. The crop yield infliction is evaluated utilizing different environmental conditions and the Regression Analysis (RA), Linear Regression (LR) Algorithm[20]

Raorane et al explains how to increase crop production by using various data mining strategies. And the strategies they used in the classification like ANN, SVM and k respectively.[21].

Kaur et al utilizes the various data mining methods to obtain higher price prediction accuracy. The data model is constructed to estimate the price. They also gathered the tomato market rate and developed the model and estimate the price through BP Neural Network and are simulated with MATLAB [22]

Gayatri et al explain the higher yield they have obtained data from multiple innovations such as IOT and web services for farmers to manage this activity and can build large volumes of data. GPS is being used to collect images from the field of agriculture and is preserved in repositories with its location.

III. METHODOLOGY

DATA COLLECTION

Data on soil nutrition were gathered from soil research in Labs. Data.gov.in offers basic crop data in the open government land. This system takes into account the primary crops including

wheat, maize, rice, jowar, and jowar, and minor crops along with pulses, jute, cotton, groundnut, barley, ragi, mustard, sugar cane, sesame.
The dataset's main characteristics are:
• Type of soil: it can be any Sandy, Clay, Silt.
• pH of Soil: This is a key measure for soil acidity and also the alkalinity.
• Soil NPK content: The three nutrients used by plants are nitrogen, potassium and phosphorus.
• Soil permeability: The soil property for water and air transmission.
• Water storage: soil water quality.
• Mean rainfall: Variations in rainfall significantly influence the supply of water to the ongoing crops from season to season.
• Temperature: It affects plant production and growth.
• Previous harvested crop: it will give support to boost the soil fertility and yield. It helps to ignore the three major issues such as soil erosion, infertility of soil and infestations of pests.
The above listed parameters of soil have a significant contribution to the ability of the crop to absorb some nutrients and also retain water. The soil nutrient quality relies on the soil pH.

Root infiltration is described by water holding ability of the soil and drainage. The composition of the soil defines the porosity of the soil and the movement of air and water on the earth so that the plants are not waterlogged.

IV. PROPOSED WORK:

Step 1: Based on common uses such as soil type, temperature, humidity, water level, size, soil ph, season, the datasets were collected and refined. Using mysql queries the collected dataset is inserted into the database. with these parameters crop name including yield rate of the crop can be estimated.

Step 2: The soil type and temperature parameters are considered as input and prediction is undertaken depends on many analyses. The soil type attribute specifies the soil type in a specific area such as Alluvial, Loamy, Black Soil, Clay and Red and the temperature attribute describes the water content in the soil.

NEURAL NETWORK

The architecture of the neural network is very common, because it can be used for several different tasks. For an ANN model like, there are some input layer, hidden layer and output layers, which is described in figure, and three layers are much needed.

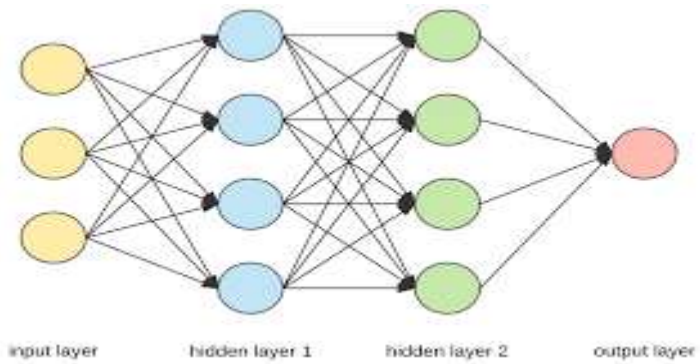

Figure 1. Neural network.

Layers such as input layer and output layer contain nodes that are related to variables of input and output, respectively. The data travels through weighted relations among layers. The node in the diagram, will get back the weights with minor error from the forward layer until the error remains constant in the network and will perform the summation of all the weights of the inputs s, defined in equation 1.

$s_k = \sum_{g=1}^{n} W_{kg} X_g$..... (1)

Here n is referred as the total inputs, W, is the weights of the related node k-g and X is defined as the input of gth node. The Transfer Function (TF) is used to determine the node output A, a sigmoid function for the hidden and output layers is the most commonly used TF, and a linear TF is usually used for the input layer.

$A_k = f(s_k)$.....(2)

The formula above seems true only for a neuron with a linear activation function (and we will get the output and it is the sum of all the weights of inputs). In a general way, it creates a non-linear, differentiable activation function.

$s_k = \emptyset(net)$ ..... (3)

Where net=$\sum_{g=1}^{n} W_{kg} X_g$..... (4)

It provides a framework for evaluating the subordinate of the error in relation to a weight using the chain law,

$\frac{\partial M}{\partial W_{kg}} = \frac{dM}{ds_k}\frac{ds_k}{dnet}\frac{dnet}{\partial W_{kg}}$ ..... (5)

Where,

$\frac{\partial M}{\partial W_{kg}}$ = error of the prediction rate changes with the weights changed.

$\frac{dM}{ds_k}$ = error of the prediction rate changes with the output changed.

$\frac{ds_k}{dnet}$= output changes with the weighted summation.

$\frac{dnet}{\partial W_{kg}}$ = weighted summation is changed with the weights changed in the data set.

EVALUATION PARAMETERS

A few of the measurement criteria are exemplified with the proposed framework for crop cost forecasting. These parameters are used to measure the efficiency of the crop price prediction. Two of the params being performed for measurement purposes are precision and also MSE. Precision is the correct measurement criterion for evaluating an efficacy of CCP.

Accuracy = $\frac{True\ positive(TP)+True\ Negative\ (TN)}{TP+TN+False\ Positive(FP)+False\ Negative(FN)}$ ..... (6)

MSE is an estimator's quality test, it's usually a positive, and the values are similar to null and superior. When y ' is a vector with n estimates, and y is also a vector with the observed input values. The basic estimation equation for the MSE is given in equation (7).

MSE = $\frac{1}{l}\sum_{k=1}^{n}(v_k - a_k)^2$ .....(7)

$(v_k - a_k)^2$ It is defined as the error square and n describes the predicted value, and v as the expected value, and an as the true value.

Here we use two hidden layers in the neural network to minimize the crop prediction error rate. The feed forward is measured in the neural network's multi-level, hidden layers.
The performance of the hidden neural network at multilevel is much higher than that of standard neural networks.

CONCLUSION

We may now agree that the crop prediction function can be performed using various methods, as mentioned above. We may also assume that ANN gives us a stronger, more detailed estimation. Thus, while using soil, climate and market rates, we will create a model with a work process, as seen above, which can provide reliable forecasts of crop yields that are appropriate for a specific region.

FUTURE ENHANCEMENT

This paper outlines the algorithm's capability to predict crop yields. In the further we will develop the appropriate algorithm based on the precision metrics, which could help to pick an effective algorithm for crop yield prediction metrics, which could help to pick an effective algorithm for crop yield prediction.

REFERENCES

[1] K. Prasanna laxmi, “A Survey on different trends in Data Streams,” 2010 International Conference on Networking and Information Technology

[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, 201