Crop Production Estimator

Mitunkumar Balar[1] Ronakkumar Patel[2]
[1,2]Department of Computer Engineering
[1,2]Charotar University of Science and Technology, India

Abstract— India is an agricultural country and agriculture plays important role in GDP of India. Due to lack of an irrigation system, the farmer mostly relies on rain. The uncertainty of rain and land per hector plays important role in a production of the crop. There are other criteria as well which affect the crop prediction like soil PH, Nitrogen, Soil Depth, temperature etc. In this paper Machine Learning approach used to predict the production of the particular crop in a particular district using rain and available land. The Neural Network is used to train the model. This will help the farmer to predict how much production he should expect and whether to go for that crop in that year. The model is evaluated using RMSE, MAE and compared with K-means clustering algorithm. Results show that the NN is giving good result compare to K-means.

Key words: Crop Prediction, Neural Network, K-means clustering

I. INTRODUCTION

Agriculture is the most important aspect in our country. The country's global development plays an important role in the large-scale economic sector. About 60% of country's land is used for agriculture to meet the necessity of 1.2 billion people. The activity of cultivation is a kind of venture at risk. Production of crop depends on various factors such as climatic, geographical, biological, political and economic factors. Agricultural precipitation and production play an important role in GDP, economic growth and national development. The Minister of Finance must also plan and manage the best storage of crops produced and be prepared to give farmers a hand in bad conditions, such as poor rainfall and heavy rain. Crop prediction is a big problem that happens. The focus of a farmer's understanding was how much he expected. Conventional farmers decide based on a past experience for specific yields, plants and weather conditions. The Character thinks directly based on the prediction of the crop that it is based on crop forecast. Accurate information on the nature of historic crop productions is an important contribution for modeling, which helps farmers and government organizations in the decision-making process to develop suitable policies for future production. Thus, modernism of cultivator is very critical and thus will lead the cultivators of our country towards benefit. There is various computer technique for prediction of crop yield or production. In this paper, Neural Network (NN) with multiple linear regression is used with deeplearning4java (DL4J) library. Artificial neural networks have proven to be a robust tool for modeling and estimation and can increase their efficiency. Regression gives various predictive accuracy like MSE (mean square error), MAE (mean absolute error), RMSE (root mean square error) and RSE (relative standard error).

II. LITERATURE SURVEY

Artificial neural networks have been exhibited to be strong tools for modeling and prophecy, to expand their efficiency.

In this study crop prediction method is used to predict the right crop by detecting different soil parameters and also atmospheric parameters. For that motive, they implemented artificial neural network (ANN) [1] and manifest the capacity of artificial neural network technology to be used for the approximation and prophecy of crop yields at countryside district.

Researcher expresses that the yield estimation is a critical problem that remains to be resolved based on handy data. Data mining techniques are the best option for this occasion. Data utilized for this proposed task are procured for the years from 2000 to 2012 for Zone of Tamil Nadu in India. The variables are Year, District, Crop, Area, Tanks, Bore Wells, Open Wells, Production and Yield [2]. This unit presents a scanty analysis of crop yield prophecy using association rule-based mining technology for the selected region i.e. district of Tamil Nadu in India. They have used k-means clustering algorithm for clustered pre-processed data and on that clustered data they applied association rule mining process. Based on that rules accuracy defined.

There are many software applications of Data Mining techniques in the field of cultivation for predicting the future year's crop production. The researchers performed Multiple Linear Regression (MLR) technique and Density based clustering technique [3] for the fixed area i.e. east Godavari district of Andhra Pradesh in India. And researcher express that

In subsequent work, a comparison of crop yield prediction can be made with all existing handy data and will be focused on suitable factors to improve the effectiveness of the proposed technique.

Author submitted project suggests a solution for smart farming by monitoring the agricultural area that can help farmers to increase productivity to a great extent. Weather condition forecast data acquired from IMD (Indian Metrological Department) [4]. This work proffers a system, in form of an android based software application, by using data analysis techniques to estimate the most fruitful crop in the current weather and soil conditions. The proposed system will consolidate the data from the repository, the meteorological service and an algorithm for machine learning: Multiple Linear Regression, the most appropriate crops are planned based on current environmental conditions.

In order to predict the crop production author used the classification strategy Support Vector Machine (SVM) and Naive Bayes [5]. AdaSVM and AdaNaive are the proposed ensemble model used to project the crop production over a period of time. This ensemble model is compared to SVM and Naive Bayes methods. The two parameters used separately for prediction of output are the accuracy and the classification error.

The Author demonstrated the crop recommendation system for maximize the crop yield. They used Ensemble technique called Majority Voting technique [6] because combining the power of two or more model gives the better prognosis, effectiveness than any of its models could attain solitary. In the voting technique, the learners they used are Support Vector Machine (SVM), NAÏVE Bayes, Multilayer perceptron and Random forest.

In the paper [7] researchers have introduced a software contrivance named 'Crop Advisor'. A web page that is ideal for estimating the effects of climatic measures on crop production.C4.5 algorithm is used to determine the climatic factors that affecting to yield of selected crops in selected districts of Madhya Pradesh. By applying the data mining technique in estimating the crop yield based on climatic input variable. A web page gives the accuracy of prediction above 75% for all the crop.

In [8] Researcher tried to solve the crop production prediction problem using k-means clustering algorithm because author express that by using k-mean algorithm they got more accurate or actual or probable prediction. Another prediction is done on the basis of rule base system using fuzzy logic by applying rule on land used for farming, rainfall and production with the help of inference engine. The researcher express that climate and precipitation are very nonlinear and complex phenomena requiring advanced computer modeling and simulation to accurately predict. Using various methodology like Back Propagation Network (BPN), Radial Basis Function Networks (RBFN), Support Vector Machine (SVM), Self-Organizing Map (SOM) they revealed that An Artificial Neural Network (ANN) [9] can be used to predict the comportment of such nonlinear systems.

In the research article [10] researcher stated that the application of neural network to the task of solving nonsequential and composite systems is optimistic. The article presents an overview of utilization of the artificial neural network to predict the crop yield using different crop performance ingredient and conclude that neural networks are one of the best solutions in search for a few agriculture problems, especially when it comes to predict crop yield.

III. PROPOSED METHODOLOGY

Neural networks are a collection of algorithms, vaguely modeled after the intelligent human brain, which are designed to recognize patterns. They interpret sensory data through a kind of perception of the machine, the labeling or cluster of unprocessed input. The patterns they recognize are digital, contained in vectors, and all the information of the real world, whether images, sounds, texts or time series, must be translated into them. The layers are made of neural *nodes*. A node is just a place where the calculus occurs, vaguely modeled on a neuron in the human brain, which triggers when it encounters enough stimuli. The node gathers input from the data with a set of coefficients, or weights that either escalate or weaken those input, thus assigning importance to inputs for the task that the algorithm is trying to absorb. (For example, which input is most useful is classifying information without fallacy?) These input weight products are collected and passed through the so-called activation function to determine whether this signal is more advanced over the network and to what extent this signal influences the final result of the classification. Figure 1 shows diagram of what single entity might look like. A node layer is a row of these neural switches that are turned on or off when the input is powered by the network. The output of each layer is at the same time the input of the next layer, from a first input layer that receives your data.

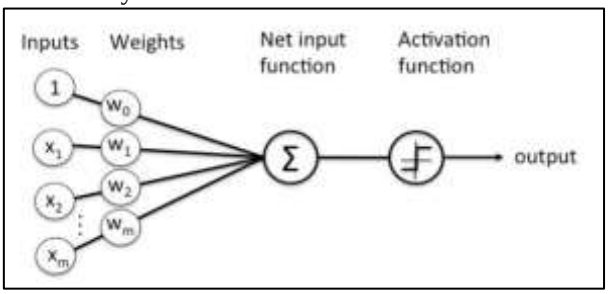

Fig. 1: Neural net with single node

Linking adjustable weights with input functions is the way we assign meaning to these functions with respect to how the network classifies and classifies the input. And an implementation time we need to care about various hyper parameters like weight initialization, epochs and iterations, learning rate, activation function, loss function, regularization, updaters and optimization algorithm, etc. Neural networks assist to cluster and classify the data. You can display them as cluster and classification layer on top of the data you reserve and organize. They help to group unlabeled data based on the matches between example entries and organize the data when they have a labeled dataset to train. (Neural networks can also extract functions that are transmitted to other algorithms for clustering and classification; so, you can consider deep neural networks as components of larger machine learning applications with algorithms reinforcement learning, Classification and Regression.). From Literature survey we found out that the implementation of NN with restricted data is missing and comparison of it with another algorithm is not done. The simple flow diagram of this model is shown in figure 2.

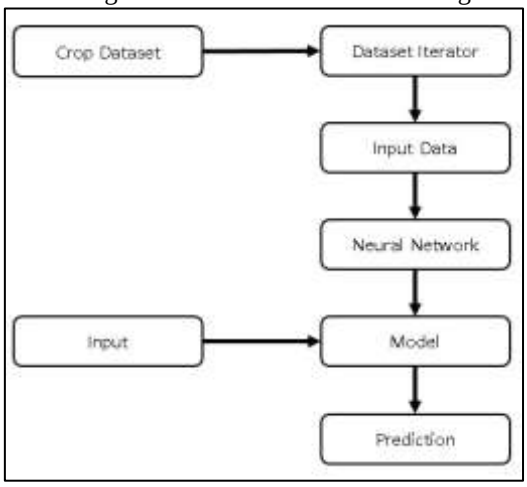

Fig. 2: Flow chart of proposed model

IV. EXPERIMENTAL SETUP

The data used in this paper is cotton data which are obtained for the year from 1997-2012 for districts Amreli of Gujarat. One data set is of crop production which is collected from the

government site (https://data.gov.in) and another data set is rainfall data which also collected from there. The features of data set are State name, District name, Crop year, Season, Crop, Annual rainfall, Area, production. Here, Area is in hector and Production is in tones. So, by combining both feature target feature decided which Production (tones/hector) it is.

A. Input:

Input file is in CSV (comma separated value) format so for reading the file we required CsvRecoredReader. And then dataset iterator will be created.
Training: After creating the iterator the training is done by neural network here we are using feed forward neural network.

B. Evaluation:

We have used here Regression method for evaluating or testing the data set. Output of the regression evaluation is in the form of various error like MAE, RMSE and RSE. Here, we have used deeplearning4java library and Intellij IDEA tool. User can also insert the data of predicted rainfall in next year from whether forecasting department in application which gives you the amount of production which you can expect in that year which is shown in figure 3.

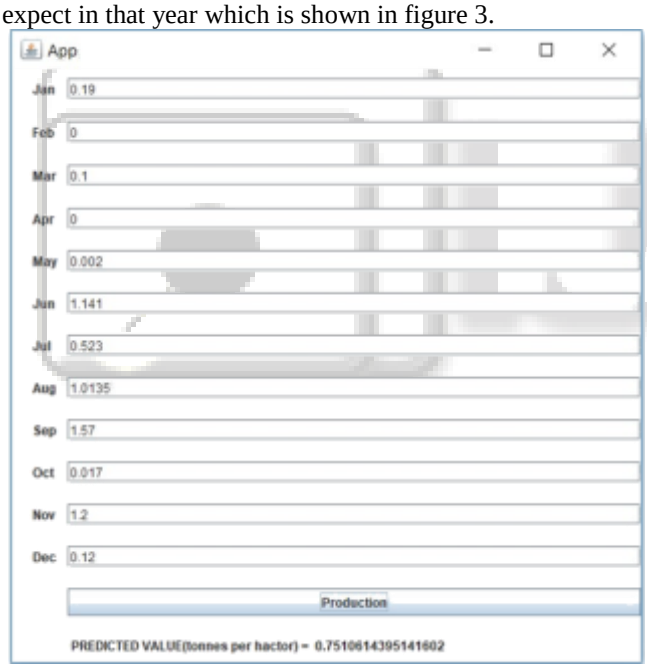

Fig. 3: GUI with predicted value

V. Result & Discussion

This article attempts to know the region-specific analysis of crop yield and it is processed by implementing both neural network and multiple linear regression. Here, neural network is used to train the model. The model is trained in respect to the past data of cotton crop of district Amreli of Gujarat.

For the training of neural net, we have used stochastic gradient descent optimization algorithm with the learning rate 0.001. The gradient descent algorithm says that the training curve should be exponentially decreased and here we also get that same curve for training. The graph is between scores and iterations. The training should be done until the score becomes constant. In figure 3 shows the training graph of predictive model.

The Predicted vs Actual production of testing data is shown in Table 1. In DeepLearning4java regression evaluation or testing is in the form of various square error i.e MAE (mean absolute error), RMSE (root mean square error) and RSE (relative standard error). Here, the output of regression evaluation of this predictive model is MAE=1.86600e-01, RMSE=2.72720e-01, RSE=1.54199e+00. In this study, the crop prediction method is used to predict the appropriate crop using only the rainfall parameter. There is some other parameter like Soil PH, Nitrogen (ppm), Depth (ppm), temperature, etc. can also be used.

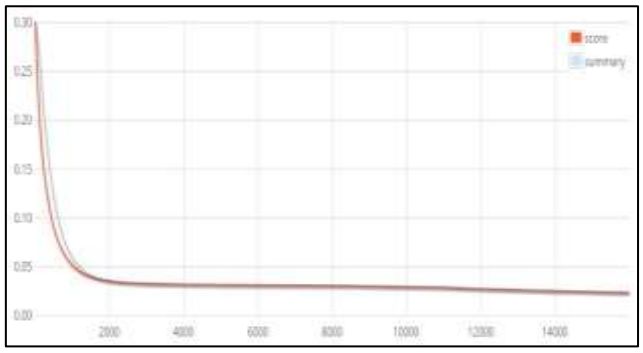

Fig. 4: Training curve of the model

Actual Production	Predicted Production
0.69	0.76
0.64	0.68
0.21	0.68

Table 1: Actual vs Predicted production of cotton

To predict the production, we have created one simple GUI in dl4j (DeepLearning4java) that provides some input fields i.e. January to December. This fields contains the rain data (in mm) of the year which we want to predict. The testing results of neural network are compared with Kmeans clustering algorithm and the results are shown in Table 2. We can interpret from the result that the NN works better than K-means.

No	Actual	Predicted		Difference	
		K-Means	NN	K-Means	NN
1	7754	9025	7968.94	1271	214.94
2	7690	9147.80	7856.86	1457.8	166.86
3	8976	9141.80	8561.36	165.8	414.64

Table 2: Comparison NN and K-means

VI. Conclusion & Future Work

In this study, we have done the prediction of cotton crop production by neural network. We have used deeplearning4java library in Intellij IDEA tool. Neural network is a very effective and reliable method that we can easily increment and decrement the feature parameters and targeted parameters. That is why we have used neural network here.

Next, we are planning to add more feature like soil PH, Nitrogen, Soil Depth, Temperature in neural network and we also want to increase the sample of the data for training purpose to increase the accuracy and effectiveness of the production. We will make GUI more user-friendly. Here we

are dealing with one crop (cotton), we want to increase the model for different crops and will recommend user which crop he needs to go in the particular season.

REFERENCES

[1] Miss.Snehal S. Dahikar, Dr. Sandeep V. Rode. Agricultural Crop Yield Prediction Using Artificial Neural Network Approach. International Journal of innovative research in electrical, electronics, instrumentation and control engineering (IJIREECE) vol.2, issue 1, January 2014.
[2] E. Manjula, S. Djodiltachoumy. A Model for Prediction of Crop Yield. International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017.
[3] D Ramesh, B Vishnu Vardhan. Analysis of Crop Yield Prediction Using Data Mining Techniques. International Journal of Research in Engineering and Technology (IJRET) vol.4, issue 1, January 2015.
[4] 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 (IJRET) vol.4, issue 5, Dec 2017.
[5] Narayanan Balakrishnan, Dr. Govindarajan Muthukumarasamy. Crop Production-Ensemble Machine Learning Model for Prediction. International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016.
[6] Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, Avinash Devare. Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique. International Research Journal of Engineering and Technology (IRJET) vol.4, issue 12, Dec 2017.
[7] S. Veenadhari, Dr. Bharat Misra, Dr. CD Singh. Machine learning approach for forecasting crop yield based on climatic parameters. International Conference on Computer Communication and Informatics (ICCCI - 2014), Jan. 03–05, 2014, Coimbatore, INDIA.
[8] Awanit Kumar, Shiv Kumar. Prediction of Production of Crops using K-mean & Fuzzy Logic. International Journal of Computer Science and Mobile Computing (IJCSMC) Vol.4 Issue.8, August- 2015, pg. 44-56.
[9] Deepak Ranjan Nayak, Amitav Mahapatra, Pranati Mishra. A Survey on Rainfall Prediction using Artificial Neural Network. International Journal of Computer Applications (0975 – 8887) Volume 72– No.16, June 2013.
[10] Siti Khairunniza-Bejo, Samihah Mustaffha and Wan Ishak Wan Ismail. Application of Artificial Neural Network in Predicting Crop Yield: A Review. Journal of Food Science and Engineering 4 (2014) 1-9.
[11] http://aps.dac.gov.in/APY/apy.csv
[12] https://download.data.world/file_download/vishuchauhan/indiandistrict-rainfall- normal-in-mm
[13] https://deeplearning4j.org

