IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-11, Nov 2018 E-ISSN: 2347-2693

Agricultural Crop Yield Prediction using Artificial Neural Network with Feed Forward Algorithm

P.Anitha1* , T. Chakravarthy2

[1]Dept. of Computer Science, A.V.V.M. Sri Pushpam College, Poondi, Thanjavur, India
[2]Dept. of Computer Science, A.V.V.M. Sri Pushpam College, Poondi, Thanjavur, India

**Corresponding Author: anithanimmi@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 17/Nov/2018, Published: 30/Nov/2018

Abstract: Rice crop production contributes to the food security of India, more than 40% to overall crop production. Variability from season to season is detrimental to the farmer's income and livelihoods. Improving the ability of farmers to predict crop productivity. In our method aimed to use of machine learning techniques Support Vector Machine (SVM), Bayesian Networks (BN) and Artificial Neural Networks (ANN) to predict rice production yield and investigate the factors affecting the rice crop yield. Data are sourced from publicly available in Indian Government's records. The attributes are used for the present studies are rainfall, minimum temperature, average temperature, maximum temperature, area, production and yield . The results showed the accuracy of, SVM is 78.76% , BN is 85.78% and ANN is 97.54% using the WEKA tool. The aim of this study are used evaluated in agriculture for predicting the crop yield production.

Keywords **–** Crop yield prediction, Crop analysis Support Vector Machine , Bayesian Networks, Artificial Neural Network.

I. INTRODUCTION

Agriculture forms the foundation of Indian economy. Large areas of agricultural land are not achieving adequate crop production due to both climatic and economic challenges. The crop production fundamentally relies upon climatic conditions, which impacts on crop development. Water shortage, absence of high yielding varieties and poor technology transfer of best agronomic practices are considered the principle factors for low crop yields in India. If crop production is to increase both crop production and the management of cropping area must improve. The use of technology and various machine learning techniques to predict crop productivity in different climatic conditions can assist farmers and other stake holders in important decision making in terms of agronomy and crop choice. Variations in climate, soil, water, pest invasion and other factors contribute to the complexity of the problem. With an increasing world population and changing climate, has come the necessity to secure the world food resources. Farmers are faced with having to make difficult decisions as to how to remain productive and sustainable with changing climates and market economic pressure. The provision of accurate and timely information such as meteorological, soil, use of fertilizers, use of pesticides can assist farmers to make the best decision for their cropping situations This could benefit them to attain greater crop productivity if the conditions are suitable or help them to decrease the loss due to unsuitable conditions for the crop yield. [1]. Agriculture is one of the important industries in the state with both food and trade products are grown in the state.[2]

The present paper is organized as follows section2 discuss about various techniques of rice yield prediction process, section3 deal with proposed method and section4 discuss about experimental results.

II. RELATED WORK

In our research work has been done to find out the efficient methods of an agriculture to predict rice yield prediction of rice production with accuracy rate.

A.K-Nearest Neighbors:

KNN method is one of the Learning Machine algorithm(LM) which is considered as a simple method to be applied in the analysis of data with many dimensions variable. This method can be used when it does not meet the classical assumptions. The performance of KNN and ensemble KNN. Although this method is simple but this method has advantages over other method. For instance, it can generalize from a relatively small training set. In This method is very important to choose the number of *k*-nearest neighbors. Ensemble technique is a method that has accuracy of prediction and efficiently used in the KNN method, so it is not necessary to search the optimal number *k*. The result shows that MAPE, MAE, and RMSE of prediction will be small if the number of *k*-nearest neighbors large. Overall, KNN ensemble method has a better

performance than KNN method. The K Nearest Neighbors (KNN) is applied for simulating daily precipitations and other weather variables [3], and different possible changes of the weather scenarios are analyzed using KNN. In KNN method is very important to choose the value of *k* -nearest neighbors number, because this can affect the predicted results. Small values of k-nearest neighbors number can produce a great variety on the prediction results, whereas a large value of k-nearest neighbor number can lead to a large biased models. KNN uses weighted average concept to parameters estimator of some single KNN models. This method is expected to yield better prediction than single KNN method.

B.Random Forest Algorithm:

RF classifier can be described as the collection of tree-structured classifiers.It is an advanced version of bagging [4] such that randomness is added to it. Instead of splitting each node using the best split among all variables, RF splits each node using the best among a subset of predictors randomly chosen at that node. A new training data set is created from the original data set with replacement. Then, a tree is grown using random feature selection. Grown trees are not pruned [5]. This strategy makes RF unexcelled in accuracy [6]. RF is also very fast, it is robust against over fitting, and it is possible to form as many trees as the user wants needs [7]. Developing accurate models for crop yield estimation using Information and Communication Technologies may help farmers and other stakeholders improve decision making in relation to national food import/exports and food security. Rice is one of the most important food crops of India. It is cultivated all over the country and contributes more than 40% of total food grain production. Given the importance of rice to world's food security, any improvements in the forecasting of rice crop yield under different climatic and cropping scenarios will be beneficial . This research has demonstrated the prediction of rice crop yield by applying one of the machine learning technique, Random Forest (RF). Only a limited number of important factors which have the highest impact on agricultural yield are selected for the random forest. The parameters selected for Rainfall ,MaximumTemperature, RiceProduction and Perception.

C.Multiple Linear Regression :

A regression model that involves more than one predictor variable is called Multiple Regression Model. Multiple Linear Regression (MLR) is the method, used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is sometimes termed as predictant and independent variables are called predictors[8]. Multiple Linear Regression (MLR) technique is based on least squares and probably the most widely used method in climatology for developing models to reconstruct climate variables from tree ring services. This crop yield prediction model is presented with the use of Multiple Linear Regression (MLR) technique where the predictant is the Production and there are seven predictors namely Year, Rainfall, Area of Sowing, Yield, Temperature, Season and Fertilizers (Nitrogen, Phosphorous and Potassium) .

III. PROPOSED WORK

In our proposed method are used three kinds of techniques such as (i) Support Vector machine(ii) Bayesian Network and (iii) Artificial Neural Network for agriculture rice yield production.

A.Support Vector Machine:

Our first approach of agriculture rice yield prediction is Support Vector Machines (SVMs).It is a supervised machine learning technique. There are a number of examples of where it has been used in the agricultural domain. This methods is applied for reduction of rainfall for climate change scenarios . To minimize the generalization error bound and to achieve generalized performance. It is used to forecast the demand and supply of pulp wood [9].It is also applied to provide insights into crop response patterns related to climate conditions by providing the features contribution analysis for agricultural yield prediction .The Sequential Minimal Optimization (SMO) classifier using the WEKA tool has been applied on the current dataset[10]. The analysis of results are made as to its effectiveness for improving rice crop yield prediction.SVM produce the accuracy rate is 78.76, but this work is produce low accuracy rate. So we select another approach called Bayesian Network.

B.Bayesian Network:

Our second approach is Bayesian Network are emerging as a valid approach for modeling and supporting decision making in other related fields such as water resource management . Another study by on crop disease has proved and also to be an effective tool[11]. The study is based on decision making of the best pest control for the scenario[12]. This model is the crop disease of condition which causes premature defoliation, weakening the plant and reducing subsequent yield. Improved prediction using BN proved to reduce the use of fungicides, producing healthier quality product and decreasing both economic costs and environmental impact[13] .

In Bayesian Network approach gives the result in quick access time but it produce the low accuracy rate 85.78%. So we select another approach called Artificial Neural Network (ANN) Feed Forward.

C.Artificial Neural Network (ANN) Feed Forward:

Our final approach is one of the powerful method in an intelligent field for predictant rice yield production. For achieving better results than both SVM and Bayesian Network. A feed forward neural network (FFNN) is the

simplest type of artificial neural network. It is consists of three layers input layer , hidden layer and output layer.[12]

The most common neural network model is the single layer perceptron that has been used in the present research[14].In this network, the information moves in only one direction, forward, from the input layer, through the hidden layers (if any) and to the output layer and the connection between the units don't form a cycle or loop [15].For present research the single layer perceptron is used as artificial neural network for the prediction of rice crop yield.

It has many advantages are produces less error, good learning ability and high accuracy rate[16]. In ANN works with WEKA tool for agriculture crop yield production is produce the accuracy rate is 97.54%.

A three layer feed-forward neural network obtained after running the technique as shown in Fig1 with 6 neurons in Input layer, 4 neurons in hidden layer and 1 neurons in output layer showing the predicted rice yield production.

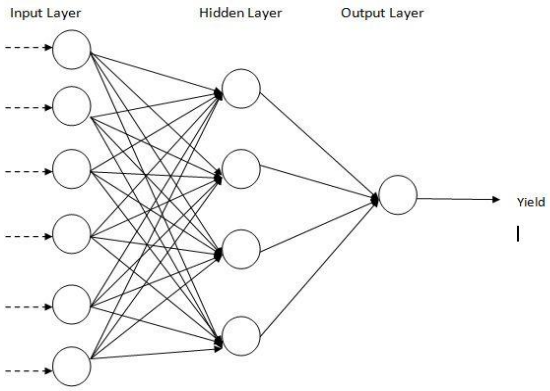

Figure1. Artificial Neural Network with Feed Forward.

IV. EXPERIMENTAL RESULTS

Perfomance Evaluation:

A confusion matrix was generated after running the current data set on Single Layer Perceptron using WEKA tool[17]. All measures can be calculated based on four values, namely True Positive (TP, a number of correctly classified that an instances positive), False Positive (FP, a number of incorrectly classified that an instance is positive), False Negative (FN, a number of incorrectly classified that an instance is negative) and True Negative (TN, a number of correctly classified that an instance is negative).

For training and testing the ANN, 10-fold cross validation method is used to subset the data. The data is randomly divided into 10 parts of which one part is used for testing while the remaining parts are used as the training data. The Feed forward algorithm is applied for training the neural network. These values are defined in Table1 .

Table 1. Confusion Matrix

		Observed	
Predicted		***True***	***False***
	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

From the values of Table1 sensitivity, specificity and accuracy are calculated.

Sensitivity = TP / (TP + FN) (1)

Specificity = TP / (TP + FP) (2)

Correctly classified instances are known as sensitivity and incorrectly classified instances are known as specificity and can be calculated using Eq. 1 and Eq. 2 described above. Accuracy is defined as the overall success rate of the classifier and computed by Eq. 3

Accuracy = (TP + TN) / (TP + FP + FN + TN) (3)

The Support Vector Machine (SVM), Bayesian Networks (BN) and Artificial Neural Networks(ANN-Feed Forward) have been used. Among the classifiers experimental results are shown in Table2.

Table 2. Comparision of classifiers performance measures

Classifier	**Accuracy**	**Sensitivity**	**Specificity**
SVM	78.76%	68.17%	83.97%.
BN	85. 78	86.31%	88.16%
ANN(FF)	97.54%	96.33%	98.12%

V. CONCLUSION AND FUTURE WORK

Our proposed method for rice crop production is three likewise. Support Vector Machine (SVM), Bayesian Networks (BN) and ANN feed forward are applied for the current study area have been proved. The ANN method gives better performance than the Support Vector Machine (SVM) and Bayesian Networks (BN).

REFERENCES

[1] R. Medar, V. Rajpurohit, "A survey on data mining techniques for crop yield prediction", International Journal of Advance Research in Computer Science and Management Studies, vol. 2, no. 9, pp. 59-64, 2014.

[2] S.Tripathi, V.V. Srinivas, R.S. Nanjundiah, "Downscaling of precipitation for climate change scenarios: a support vector

machine approach", Journal of Hydrology, vol. 330, no. 3, pp.621-640, 2006.
[3] Zhou HZ., "Ensemble Methods Foundations and Algoritms" Florida(US): CRC Pr, 2012.
[4] Archer K. J., Kimes R. V.," Empirical characterization of random forest variable importance measures", Computational Statistics & Data Analysis,52(4), 2249-2260,2008.
[5] Bureau A., Dupuis J., Falls K., Lunetta K.L., Hayward B., Keith T.P., Van Eerdewegh P.,"Identifying SNPs predictive of phenotype using random forests", Genetic epidemiology, 28(2), 171-182, 2005.
[6] Horning N., "Random Forests: An algorithm for image classification and generation of continuous fields data sets", New York,2010.
[7] Liaw A., Wiener M., "Classification and regression by random Forest", R news, 2(3),18-22, 2002.
[8]G.Ruß, "Data Mining of Agricultural Yield Data : A Comparison of Regression Models", Conference Proceedings, Advances in Data Mining – Applications and Theoretical Aspects, P Perner (Ed.), Lecture Notes in Artificial Intelligence 6171, Berlin, Heidelberg, Springer, pages : 24-37,2009.
[9] S. Brdar, D. Culibrk, B.Marinkovic, J.Crnobarac , V. Cmojevic, "Support Vector Machines with Features Contribution Analysis for Agricultural Yield Prediction", 2011.
[10]M.Trnka, "Projections of Uncertainties in Climate Change Scenarios into Expected Winter Wheat Yields", Theoretical and Applied Climatology, vol. 77, pages : 229-249, 2004.
[11] Mehta D R., Kalola A D., Saradava D A.,Yusufzai A S., "Rainfall Variability Analysis and Its Impact on Crop Productivity - A Case Study", Indian Journal of Agricultural Research, Volume 36, Issue 1, pages : 29-33, 2002.
[12] A. Castelletti, R. Soncini-Sessa, "Bayesian Networks and participatory modelling in water resource management", Environmental Modelling & Software, Vol. 22, No.8, pp.1075-1088, 2007.
[13] C. Bi, G. Chen, "Bayesian networks modeling for Crop Diseases", Computer and Computing Technologies in Agriculture IV, Springer Berlin Heidelberg, pp. 312-320, 2010.
[14] Auer, Peter, Harald Burgsteiner, Wolfgang Maass , "A learning rule for very simple universal approximators consisting of a single layer of perceptrons" (PDF). Neural Networks. **21** (5): 786–795. doi:10.1016/j.neunet.2007.12.036. PMID 18249524,2008.
[15] A. Abraham, "Artificial Neural Networks: Handbook of Measuring System Design", John Wiley & Sons, Ltd., 2005,ISBN: 0-470-02143-8.
[16]https://en.wikipedia.org/wiki/Types_of_artificial_neural_networks.
[17] Weka 3:Data Mining Software in Java, Machine Learning Group at the University of Waikato, Official Web: http://www.cs.waikato.ac.nz/ml/weka/index.html, accessed on 26th March 2016.

