International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 1 pp. 368-371 Jan 2018-June 2018 www.csjournals.com

Yield Prediction using Machine Learning: A Review

[1]Sunila Godara,[1]Sanjeev Kumar
[1,2]Department of CSE, Guru Jambheshwar University of Science & Technology, Hisar, India

Abstract: Now-a-day's agriculture is one of the most important field and backbone of our country. Poor condition of agriculture is due to not well defined guidance to the farmers. This improper farming affects the yield of crop. If the farmer is attentive about the crop cultivation techniques and harvesting it will extra supportive for farmers to maximize the crop productivity. Machine Learning (ML) can be used for this purpose. Crop yield estimation can be done using different ML techniques like K-Means, K-Nearest neighbor (KNN), Artificial Neural Networks(ANN) and SVM. This paper presents a brief review of various ML techniques used for crop yield prediction.

I.Introduction

According to statistics of 2016 approximately 272.82 million farmers dwell in Maharashtra. Increasing suicide of farmers there is prim need to recognize the significance of prior crop prediction. Basic knowledge about soil and location-wise weather conditions helps to achieve high crop yield through technology [16],[2]. Yield prediction is aextremely significant agricultural dilemma. Any farmer is interested in knowing how much yield he is about to produce. In the ancient times, yield prediction was done by farmer'sknowledge of specific field.In this situation datawere available from some past time period, where the equivalent yield predictions have been done. Fig1 shows variables and Framework for yield prediction.

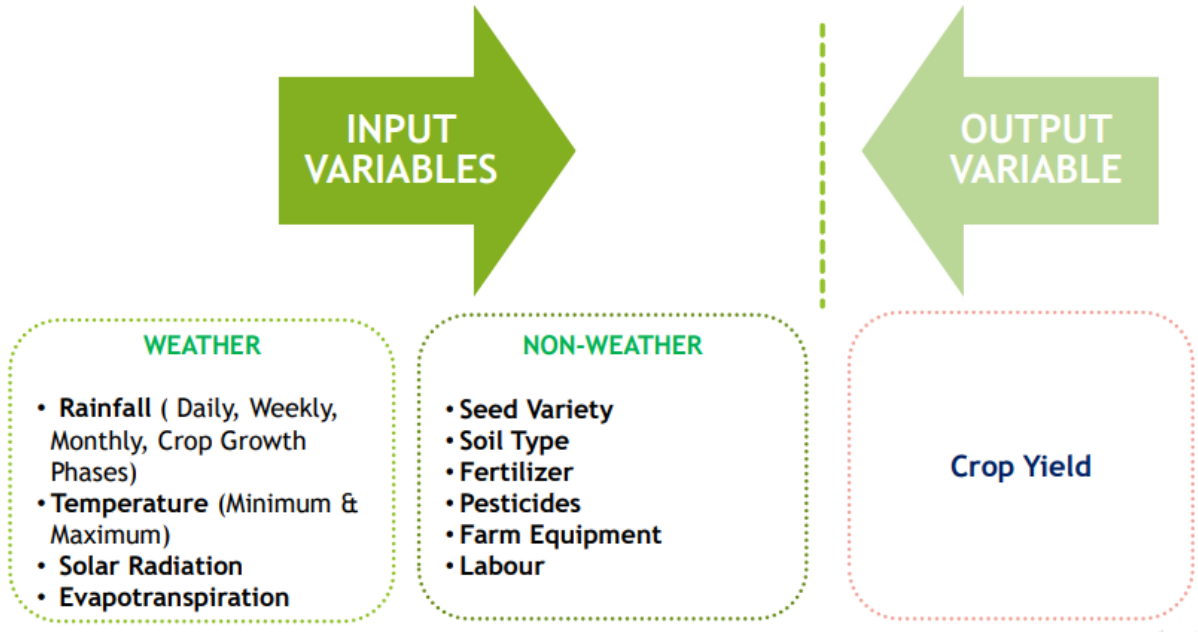

Fig1: Variables and Framework for yieldprediction [15]

ML techniques can be used to predict future crop, weather forecasting, pesticides and fertilizers to be used, revenue to be generated and so on. Machine Learning (ML)techniques can be supervised or unsupervised.

If the training data is to be collected from some time back and used for training which has to be exploited to learn how to classify future yield predictions is known as supervised ML. Data Mining techniques were discussed by the researchers [2]. Multilayered Perceptron, Radial Basis Function,

Support Vector Machine and Decision Tree were analyzed. Their comparison showed that the Support Vector Machine technique is the most suitable for Yield prediction [3].

Clustering is widely used as unsupervised technique for yield prediction. A sub-set of object which is similar in nature called a cluster. Highly similarobjects are placed in same clusters and the objects are dissimilar in other clusters. The different clustering methods are Hierarchical Methods(HM), Partitioning Methods (PM), Density-based Methods(DBM), K-Meanand Grid-based Methods. The K-Means is the most important clustering technique. This technique is used to group the data which have no previous information about the data. The value of K tells the number of clusters in which data is divided. For K numberis used to define K centers. These centers placed at a distance from each other and then associate each data to the cluster that has the closest centroid. The process of finding new K centers and assigning data sample to the clusters that has the closest centroid is iteratively carried out until no longer the data samples can change their clusters. Using this K-Means approach the Government could increase their production practices by acquiring new farmers or framing new profitable agriculture schemes based on crop yield prediction [2]. Weighted Class Based Clustering (WCBC)[1] inspired from K-means algorithm enhanceseprability of data objects on the basis of proposed weighted distance measure in which weights are calculated from class labels and class labels further contains classifying properties of data itself[1].

II. Literature Review

This section gives a review of various papers that is figure out the crop yield production using various ML techniques.

RatchaphumJaikla et al. [4] developed rice yield prediction method using the Support Vector Regression method (SVR). In this Paper yield prediction method is divided into three phases, i.e., soil nitrogen prediction, rice stem weight prediction and rice grain weight prediction. The results were compared using DSSAT4 with help of Crop Simulation Model (CSM-Rice simulation model). The results obtained were comparable with CSMRice simulation model.

Wang peijuanet al.[5]developed,RS-P-YEC (Remote Sensing – Photosynthesis –for Crop) a radiative transfer equation for winter wheat Yield Estimation. This model usedharvest index to convert the net primary productivity to wheat yield . It is helpful in the yield assessment of wheat in North China withextensivelyaccessible remotely sensed images. The resultsshowed that there wasbig variations in yield of wheat transversely in North China . In the southern region, the yield is uppermost with supplementary solar radiations.

Jianqiang REN et al.[6] studied the wheat yield assessment using retrieved Leaf Area Index (LAI) from remote sensing in North China. S-G filter was used to improve the quality of data and also decrease the error of yield estimation. Gaussian model was also used at the same time to simulate daily crop LAI..The relationship wasmade between LAI and yield of crop. After optimization of yield estimation model, the crop period and most excellent model wereselected. LAI from MODIS-NDVI was taken forestimation winter wheat yield. In the results average relative error was low at 1.21% and that RMSE was 257.33 kg ha-1.

Zongnan Li et al. [7]usedremote sensing for monitoring the crop growth at different scales.The canopy spectra and leaf area index (LAI) of dissimilar crop growth conditions were collected and analyzed. Several techniqueshave been prepared toreduce the effects of the soil and non-vegetation fraction. These techniques were new vegetation indices (VIs), soil-adjusted vegetation index (SAVI), Modified Soil Adjusted Vegetation Index (MSAVI) and Enhanced Vegetation Index (EVI).HJ-1A satellite's multi-spectral remote sensing data were used to learn the crop development monitoring indicators.At early stage best indicator was SAVI (L=0.3). In combining stage, the best indicator was SAVI (L=0.2). The results further showed that the vegetation indices used to decrease soil-effect are appropriate for the area in dissimilar crop coverage in early stage .To pick up precision of regional crop, extradata related tocrop coverage is needed. With help of this crop coverage information, VI and optimal parameters can be decided for different stages.

BoonyasithKhobkhun et al. [8]presented a method to establish rice cropping pattern in Thailand for prospect prediction of water supply demand, pricing and extra issues as well as governmental policies. Moderate-Resolution Imaging SpectroraDiometer(MODIS) operated by NASA was used to accumulate data. ANormalized Difference Vegetation Index (NDVI) wasobtained from MODIS

datasets once after every 16 days. A Progressive Iteration Approximation(PIA) was usedforsignalsmoothingand reducing noise. The results showed that using PIA technique for noise decline has provided better results in comparison ofordinary filtering technique like SavitzkyGolay filter.

AakunuriManjulaet al.[9]proposed extensible crop yield prediction framework (XCYPF) that is bendable and extensible. It gives provision for choice of crop,selectionsof dependent and independent variables, and selections of datasetsfor crop yield prediction. Variety ofindex like temperature condition index TCI, vegetation condition index VCI and normalized difference vegetation index NDVI wereused for accessing productivity of crops.

Leisa J. Armstrong et al. [10] examineddata visualization techniques to determine correlations between the climatic factors and rice crop yield. Data mining techniques were usefulto predict rice crop yield for Kharif season of Tropical Wet and Dry climatic zone of India using agriculture dataset.WEKA tool was used for classification in this paper. Research shows thatas decrease in precipitation in the selected climatic zone that will increase the rice crop yield. J48 and LADTree had achieved maximum accuracy.

Niketa Gandhi et al.[11]proposed a modified decision support System for yield prediction of rice crop in Maharashtra state. The system allows the selection of the range of precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapotranspiration and predicts the expected class of yield viz., low, moderate or high. Past data is used to calculate parameters. The three classes for the yield prediction were defined as low, moderateand high.

Leisa J. Armstrong et al. [12]in this research work rice yieldproduction was done with neural networks.Precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapotranspiration, area, production were taken as parameters for yield the Kharif season for the years 1998 to 2002. A Multilayer Perceptron Neural Network was developed using WEKA.Validation of results was done using cross validation method. The results showed the high accuracy up to 97.5% with a sensitivity of 96.3 and specificity of 98.1.

Pritam Bose et al.[13]used spiking neural networks (SNNs) for remote sensing spatiotemporal analysis of image time series for crop yield judgment. MichielStas, Jos Van et al.(2016)used Boosted Regression Trees (BRT) and Support Vector Machines (SVM) for yield prediction. Single NDVI, Incremental NDVI,Targeted NDVI. BRT and SVM were first used to select features with elevatedsignificance for prediction of yield. After feature selection, BRT and SVM models were compared to the subset of selected features for yield forecasting. BRT outperformed SVM.Fig2 shows various Machine Learning Techniques used for yield prediction[15].

CONCLUSION

Agriculture is strongly participating in the GDP of India. According to the latest statistics agriculture has 17.32% participation in the GDP of India. So a correct measure in the yield prediction of crops before harvesting is crucial need of the country.Use of ML in agriculture can revolutionize the picture of decision making and farmers can yield in better way. Many models and techniques proposed by the researchers has been studied during the literature survey of yield prediction of crops. This paper provides the work of various authors to get information of current scenario of ML techniques and applications in yield prediction.

REFERENCES

1. Sunila, Rishipal Singh and Sanjeev Kumar. "A Novel Weighted Class based Clustering for Medical Diagnostic Interface." Indian Journal of Science and Technology Vol 9, Issue 44, Nov.2016.
2. A.Senthil Kumar, P.Arun., “A Survey on Agriculture Analysis for Crop Yield Prediction Using Data Mining Techniques”, IOSR Journal of Computer Engineering (IOSR-JCE), pp. 23-25, 2001.
3. 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, 24–37, Springer, 2009

4. Jaikla, R., Auephanwiriyakul, S., Jintrawet, A., "Rice Yield Prediction using a Support Vector Regression method", IEEE proceedings of ECTI-CON, pp, 29-32, 2008.
5. Peijuan,W.,Jiahua, Z., Donghui, X., Yuyu, Z., Rui, S., "Yield estimation of winter wheat in north china plain using RS-P-YEC model", IEEE International Geoscience and Remote Sensing Symposium, pp, 378-381, 2009.
6. Ren, J., Chen, Z., Yang, X., Liu, X., Zhou, Q., "Regional yield prediction of winter wheat based on retrieval of leaf area index by remote sensing technology", IEEE International Geoscience and Remote Sensing Symposium, pp, 374-377, 2009.
7. Li, Z., Chen, Z., "Remote sensing indicators for crop growth monitoring at different scales", IEEE International Geoscience and Remote Sensing Symposium, pp, 4062-4065, 2011.
8. Khobkhun, B., Prayote, A., Rakwatin, P., Dejdumrong, N., "Rice phenology monitoring using PIA time series MODIS imagery", IEEE 10th International Conference Computer Graphics, Imaging and Visualization, pp, 84, 2013.
9. Manjula, A., Narsimha, G., "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", IEEE Sponsored 9th International Conference on Intelligent Systems and Control, 2015.
10. Gandhi, N.,Armstrong, L.J., "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques", IEEE International Conference on Advances in Computer Applications, pp, 357-363, 2016.
11. Gandhi, N.,Armstrong, L.J., Petkar, O., "Proposed Decision Support System (DSS) for Indian Rice Crop Yield Prediction", IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development, pp, 13-18, 2016.
12. Gandhi, N.,Armstrong, L.J., Petkar, O., "Rice Crop Yield Prediction Using Artificial Neural Networks", IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development, pp, 105-110, 2016.
13. Bose, P., Kasabov, N.K., "Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series", IEEE transactions on geoscience and remotesensing, vol. 54, no. 11, pp, 6563-6573, 2016.
14. Stas, M., Orshovn, J.V., Dong, Q., Heremans, S., Zhang, B., "A Comparison of machine learning algorithm for regional wheat yield prediction using NDVI time series of SOPT-VGT", IEEE International Conference Agro-Geoinformatics, 2016.