INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 3)
Available online at: www.ijariit.com

Rice crop yield prediction using machine learning techniques

Suhas L.
hegdesuhas27@gmail.com
SJB Institute of Technology, Bangalore, Karnataka

Sangamesh
sangameshg08@gmail.com
SJB Institute of Technology, Bangalore, Karnataka

Prakash kumar
prksjbit1197@gmail.com
SJB Institute of Technology, Bangalore, Karnataka

Supriya B. N.
bnsupriya@gmail.com
SJB Institute of Technology, Bangalore, Karnataka

ABSTRACT

In an agricultural country like India where agriculture is the main source of occupation for most of the people, there are many factors that contribute to the total yield of the crop being cultivated. Depending on the various soil conditions and climatic conditions the yield of the crop might vary. It would be helpful for the farmer to know which crop gives more profit in certain climatic condition. This paper proposes the use of a machine learning technique to help farmer predict the yield for a particular year.

Keywords— *Machine Learning, Multilinear regression, Decision tree regression, Crop prediction*

1. INTRODUCTION

In India agriculture is the main source of income for most people. Agriculture is a very important employment venture of the most family in India. Rice is an imperative product and used broadly over the globe as staple sustenance. India is among the main rice makers on the planet. India produces rice in an extensive amount and in the last money related year 2014-2015, rice creation was in excess of 100 million tons. Aside from the main maker, India is additionally one of the biggest exporters of rice. In the last budgetary year, India sent out in excess of 8 million tons of rice to numerous nations. Given the significance of rice to the world's sustenance security, any enhancements in the gauging of rice edit yield under various climatic and trimming situations will be advantageous.

Rice crop yield for a particular year depends on the water level, soil type, mean temperature, maximum and minimum temperature etc. In this paper, we propose a system that predicts the rice crop yield for a particular year for a particular area using the regression techniques. We use a multiple linear regression model and the decision tree regression model to predict the crop yield. Various parameters like water Evo transpiration, temperature, previous year yield etc. are used to train the machine learning model. Once the model is trained, it will predict the rice crop yield of a particular year based on the data we input for that year. This prediction will help the farmer to estimate the approximate yield he will get in that particular year even before he starts the process of cultivating the crop.

2. RELATED WORK

Heamin Lee et al. [7] designed an IoT system for disease and pest prediction in the orchard. All the weather conditions affect the growth of plants. Shufen Zhang et al.[8] proposed a system which uses ZigBee network to connect the terminal sensing devices, and connect the big data platform by IoT designed for wheat diseases, pests and weed with the expert system. Abdullah Na et al. [9] designed a system for remote monitoring of soil characteristics. For a farmer, it is important to have knowledge about soil and its characteristics so based on it he can develop strategies. Studies have uncovered the use_of_back propagation framework to_predict_rice_yield in light of atmosphere data [13]. Conjecture-of-maize yield in light of precipitation, soil and diverse parameters got a trying misstep of_14.8% [14], and expectation of rice yield utilizing climatic perception information anticipated with the greatest of_45-60kg/ha [15]. Distinctive examinations have utilized neural frameworks to envision rice yield in perspective of soil parameters and achieved a testing oversight of 17.3% [16] or developed an immediate and correct estimation to anticipate rice yields [17]. Item yield figure utilizing ANN was also broke down for_Nepal.

Kogan et al. [9] compared different methods for winter wheat yield forecasting: using remote sensing observations, meteorological data and biophysical models.

3. IMPLEMENTATION

We use Multi Linear Regression (MLR) and Decision Tree Regression (DTR) to predict the rice crop. Various steps used in this process are,

- Data acquisition
- Data Pre-processing.
- Building and train machine learning model
- Prediction of crop yield
- Compare model performance.

© 2019, www.IJARIIT.com All Rights Reserved

We first collect the various data required to build the prediction model from the Indian Government Websites. The data collected are not in the ready to use format. Some data might be missing or unavailable. We deal with such exceptions in the form of data preprocessing using the normalization techniques. Once the data is normalized, we use it to build the machine learning models. The processed data is split into training and testing data. We use 80% 0f the data to train the system and 20% of data to test the system that is built. Once the model is built, we feed the system with the new dataset of a future year for which the prediction is to be done.

Once the prediction for the new set of data is done, we calculate the Mean Square Error (MSE) and Co-efficient of Determination (CoD) for the MLR and DTR models. MSE and CoD help us to identify the accuracy of the two models.

If a vector of n predictions generated from a sample of *n* data points on all variables, and Y is the vector of observed values of the variable being predicted, then the within-sample MSE of the predictor is computed as

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

R^2 (or r^2), a measure that assesses the ability of a model to predict or explain an outcome in the decision tree regression setting. More specifically, R^2 indicates the proportion of the variance in the dependent variable (Y) that is predicted or explained by linear regression and the predictor variable (X, also known as the independent variable).

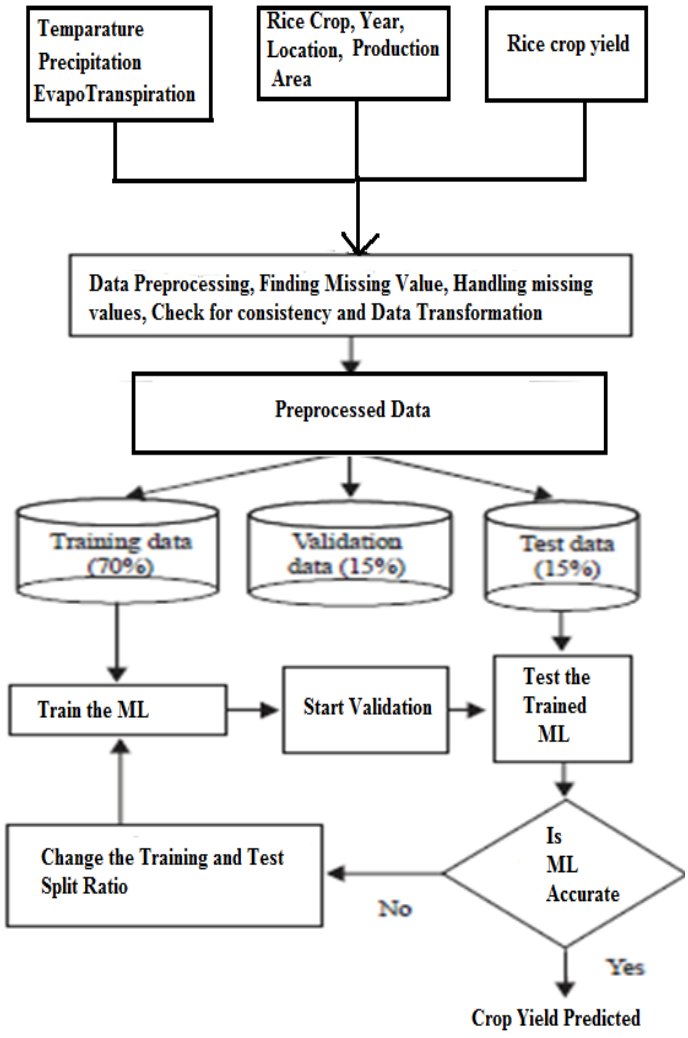

Fig. 1: Proposed system

4. RESULTS AND CONCLUSION

In this chapter results obtained from the proposed algorithms are discussed. First, the standard agriculture dataset available in the Indian government web sources is collected. The various parameters such as precipitation, temperature, reference crop evapotranspiration and production are analyzed. These parameters are the independent variable and crop yield is the dependent variable, which needs to be evaluated. Once the crop yield is predicted we measure the accuracy of the two models using the MSE and CoD.

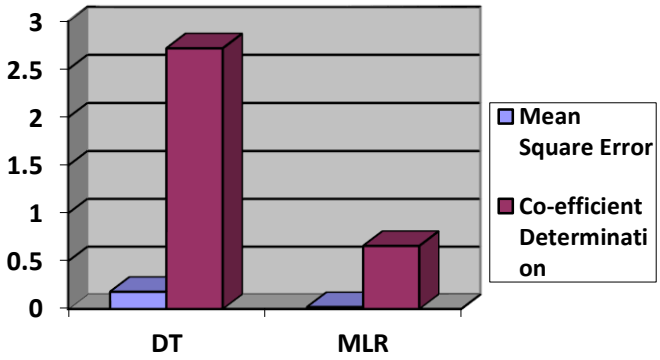

Fig. 2: MSE and Co-efficient Determination values of DT and MLR

Table 1: Accuracy results obtained of DT and MLR

Measures	DT	MLR
Mean Square Error	0.18	0.02
Co-efficient Determination	2.72	0.66

5. FUTURE WORK AND CONCLUSION

Application of machine learning is proposed in this project to help farmers and agricultural industry for making the decision better in crop yield. Crop data on rice of various district is collected from the web resource is collected for exploratory data analysis. Explored decision tree regression and multi-linear regression machine learning techniques can compare accuracy result for different agriculture crop data. The result obtained in this proposed project can help farmers to know and understand the crop yield prediction on various crops with different depended parameters such as temperature, waterfall, and evapotranspiration and so on.

As of now, we are foreseeing just the yield for particular products. We can include the forecast for venture required for various yields. Components like manure, pesticide, arrangement of the ranch for sowing, cultivate gear and bore wells assume a vital part in choosing which yield to develop. We can likewise propose compost supplement needs of the dirt if the agriculturist gives the dirt examination comes about. The recommendations for intercropping and money harvests can likewise be added to expand profitability. A few vegetation files like Normalized Difference Vegetation Index (NDVI), Vegetation Condition Index (VCI) and Temperature Condition Index (TCI) can be utilized to distinguish draft conditions and a few other climate impacts on the yield of the harvest.

6. REFERENCES

[1] Y.A. Pachepsky, D. Timlin, and G. Várallyay, "Artificial neural networks to estimate soil water retention from easily measurable data", Soil Science Society of America Journal, vol. 60(3), pp.727-733, 1996.

[2] D.A. Elizondo, R.W. McClendon and G. Hoogenboom, "Neural network models for predicting flowering and

© 2019, www.IJARIIT.com All Rights Reserved

physiological maturity of soybean", Transactions of the ASAE, vol. 37(3), pp.981-988, 1994.

[3] D.A. Elizondo, R.W. McClendon, and G. Hoogenboom, "Neural network models for predicting flowering and physiological maturity of soybean", Transactions of the ASAE, vol. 37, pp. 981–988, 1994.

[4] C.C. Yang, S.O. Prasher, S. Sreekanth, N.K. Patni and L. Masse, "An artificial neural network model for simulating pesticide concentrations in soil", Transactions of the ASAE, vol. 40, pp. 1285–1294, 1997.

[5] M.G. Schaap and W. Bouten, "Modeling water retention curves of sandy soils using neural networks", Water Resources Research, vol. 32, pp. 3033–3040, 1996.

[6] W.D. Batchelor, X.B. Yang and A.T. Tshanz, "Development of a neural network for soybean rust epidemics", Transactions of the ASAE, vol. 40, pp. 247–252, 1997.

[7] G.A. Thomas, G. Taylor and J.C.Wood, "Mapping yield potential with remote sensing", Precision Agriculture, vol. 1, pp.713–720, 1997.

[8] M. Kaul, R.L. Hill and C. Walthall, "Artificial neural network for corn and soybean prediction", Agricultural System, vol. 85, pp. 1-18, 2005

[9] B. Ji, "Artificial neural networks for rice yield prediction in mountainous regions", Journal of Agricultural Science, vol. 145, pp. 249–261, 2007.

[10] B.A. Smith, G. Hoogenboom and R.W. McClendon, "Artificial Neural Networks for Automated Year round Temperature Prediction", Computers and Electronics in Agriculture, vol. 68, pp. 52–6, 2009.

[11] M. Schaap and W. Bouten, "Modeling water retention curves of sandy soils using neural networks", Water Resour.Res., vol.32, pp. 3033– 3040, 1996.

[12] S.K. Starrett and G.L. Adams, "Using artificial neural networks and regression to predict the percentage of applied nitrogen leached under turfgrass", Commun.Soil Sci. Plant Anal., vol. 28, pp.497–507, 1997.

© 2019, www.IJARIIT.com All Rights Reserved