International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

ADVANCED CROP YIELD PREDICTION USING REGRESSION ANALYSIS

M.Sundhari [1], Varsha.P.B [2], Divyabharathy.R [3], Harini.T[4]

[1]Assistant Professor, [2],[3],[4]UG SCHOLAR

Department of Information Technology, Panimalar Institute of Technology, Chennai

msundhari06@gmail.com [1], varshasri073@gmail.com [2],bharathidhivya053@gmail.com [3], harinitamil30@gmail.com [4]

Abstract - *Crop production contributes to the food security of India,morethan40% to overall crop production. This paper describes the crop prediction system which can predict the yield of crop using various factors. Thus ,we are implementing a method by using technology to predict the crop yield under any sort of situation where they outcome depends on the particular input at any time T. We accomplish this by using machine learning and regression analysis for predicting the crop yield with k-fold techniques and batch training. We make sure to predict a clear success rate of the crop yield so the decisions can be made clearer and more accurate as much as possible.*

Keywords- Regression Analysis, decision tree, Machine Learning ,crop prediction, k-fold.

1. INTRODUCTION

Accomplishing high harvest yields is the guideline point of rural creation. The acknowledgment and the board of variables that impact crop yield help ranchers in basic leadership. There is various harvest yield forecast models which utilize either measurable or crop reenactment models. Throughout the most recent decade it has been seen that Machine Learning strategies give an increasingly successful way to deal with foreseeing crop yield under various trimming situations .The utilization of Machine Learning can make models with complex sources of info less demanding to decipher. This exploration portrays the advancement of a rice crop yield forecast model using Machine Learning. The forecast is accomplished with the adaptable approach of Machine Learning utilizing back proliferation system. The most well-known AI display is the multilayer perceptron that has been utilized in the present research. This methodology has been exhibited by gauging of rice crop yield expectationforKharifseasonfromyear1998to2002forMaharashtraprovince of India, based on various indicator factors including precipitation, least temperature, normal temperature, most extreme temperature, reference crop evapotranspiration and yield. AI with Multilayer Perceptron were considered for the presenter search.

II. RELATED WORKS

[1] Artificial neural networks for rice yield prediction in mountainous regions

This system aims at (1) investigate whether artificial neural network (ANN) models could effectively predict Fujian rice yield for typical climatic conditions of the mountainous region, (2) evaluate ANN model performance relative to variations of developmental parameters and (3) compare the effectiveness of multiple linear regression models with ANN models. Models were developed using historical yield data at multiple locations throughout Fujian.

[2] Neural network prediction of maize yield using alternative data coding algorithms

This system uses Back propagation neural networks with five data coding schemes to predict maize yield at three scales in east-central Indiana of the Midwest USA, using 1901–1996 local crop-stage weather data and yield data from farm, county, and state levels. Input data included precipitation and air temperature during maize reproductive (R) stages R1 (silking) to R5 (denting of kernels), the year, and, for some nets, the scale of yield data.

[3] Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand

This study aims to develop the decision support system using Artificial Neural Networks (ANN) by adjust the value of parameters and study about 9 Algorithms training. In predicting rice productions which its study found that each values that was adjusted making high predicting like appropriate number of hidden nodes to model equals to 9, learning rate effects the speed of appropriate learning to the model equals to 0.5, and appropriate momentum to model was 0.5.

[4] D. Back propagation algorithm for rice yield prediction

In this study, a Back propogation (BP) algorithm is utilized to develop a neural network model to predict rice yield based on the aforementioned factors in MUDA irrigation area, Malaysia. The result of this study shows that the BP algorithm is able to predict the rice yield to a deviation of less than 0.21.

PROPOSEDSYSTEM

This study aimed to use neural networks to predict rice production yield and investigate the factors affecting the rice crop yield for various districts of Maharashtra state in India. Data were sourced from publicly available Indian Government's records for 27 districts of Maharashtra state, India. The parameters considered for the present study were precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapo transpiration, area, production and yield for the Kharif season (June to November) for the years 1998 to 2002. The dataset was processed using WEKA tool. A Multilayer Perceptron Neural Network was developed. Cross validation method was used to validate the data. The results showed the accuracy of 97.5% with a sensitivity of 96.3 and specificity of 98.1.

IMPLEMENTATION MODULES

A. PRE-PROCESSING:

Pre-processing refers to the transformations applied to our data before feeding it to the algorithmic program. Data Pre-processing is also a way that is accustomed convert the {raw data information information} into a clean information set. In different words, whenever the information is gathered from utterly completely different sources it's collected in raw format that may not potential for the analysis.

B. DATACLEANING:

Data cleaning is to spot and take a way errors& duplicate knowledge, so as to make a reliable dataset. This improves the standard for the coaching knowledge for analytics most knowledge scientists pay a massive quantity of your time enhancing the standard of the information.

C. DATAINTEGRATION:

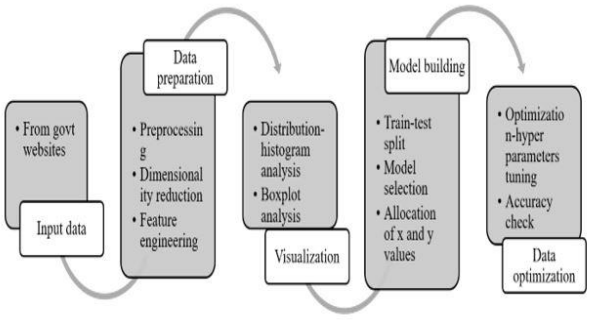
Data integration involves combining information residing in several sources and providing users with a unified read of them. This method becomes important during a type of things ,that embrace ach business and scientific domains.

D. DATATRANSORMATION:

Data transformation is simple or sophisticated supported the desired changes to info the info the information} between the availability (initial) information and thus the target (final) knowledge.

E.DATAVISUALIZATION:

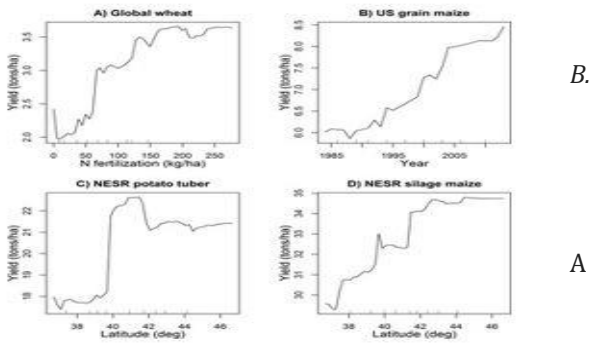
Data visualization could be a crucial step for building a strong and economical machine learning model. It helps U.S.A. to raised perceive the information, generate higher insights for feature engineering, and, finally, build higher choices throughout modeling and coaching of the model.



Architecture diagram

ALGORITHMS

Performance of RF for global and regional crop yield predictions. Accurate predictions of crop yield are critical for developing effective agricultural and food policies at the regional and global scales. We evaluated a machine-learning method, Random Forests (RF), for its ability to predict crop yield responses to climate and biophysical variables at global and regional scales in wheat, maize, and potato in comparison with multiple linear regressions (MLR) serving as a benchmark.

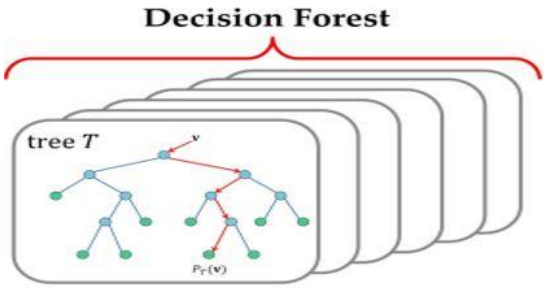


B.

A

DECISIONTREE

Decision tree is a decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. As the name goes, it uses a tree-like model of decisions. Though a commonly used tool in data mining for deriving a strategy to reach a particular goal, its also widely used in machine learning, which will be the main focus of this article.



Decision Tree

RESULT

This will help us to understand the increase/decrease in production for a particular region and help in determining the proactive measures for the results. Accomplishing high harvest yields is the guideline point of rural creation. The acknowledgment and the board of variables that impact crop yield help ranchers in basic leadership. There is various harvest yield forecast models which utilize either measurable or crop reenactment models. Throughout the most recent decade it has been seen that Machine Learning strategies give an increasingly successful way to deal with foreseeing crop yield under various trimming situations.

CONCLUSION

From this prediction it is evident that our model performs well, and our overall accuracy is 95%.This supports our claim about our objective. After prediction , back propagation can be implemented to understand many insights like -Why crop production has decreased in a particular region, root cause can be identified and proactive measures can be implemented in that particular region.

FUTURE SCOPE

With provided more data across multiple regions across will help the model to converge and forecast the crop protection index i.e., model can be applied an any regional-bound. Feature selection with other dependencies with affects the agriculture in terms of their protection according to their specific region can be added for more accurate predictions. Many comparisons of ML methods for yield prediction have been made, seeking for the most accurate technique.

REFERENCE

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

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

[3] 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.

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

[5] J. Liu, C. Goering and L. Tian, “A neural network for setting target corn yields”, Transaction of the ASAE, 44(3), pp. 705-713, 2001.

[6] M. O‘Neal, B. Engel, D. Ess, J. Frankenberger, “Neural network prediction of maize yield using alternative data coding algorithms”, Biosystems Engineering, 83(1), pp. 31-45, 2002.

[7] S. Puteh, M. Rizon, M. Juhari, J. Nor Khairah, S. Siti Kamarudin, B. Aryati, R. Nursalasawati, “Back propagation algorithm for rice yield prediction”, Proceedings of 9th of the Ninth International Symposium on Artificial Life and Robotics, Beppu, Japan, Oita, pp. 586-589, 2004.

[8] W. Ji and J. Cui, “Application of Geographical Information System (GIS) in Agricultural Land classification and Grading”, 2011 International Conference on Agricultural and Natural Resources andEngineering Advances in Biomedical Engineering, 3(5), pp. 201-205,2011.

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

[10] T. Ranjeet and L. Armstrong, “An architecture of a decision support system for Western Australian Agriculture Industry”, Proceedings of the 9th Conference of the Asian Federation for Information Technology in Agriculture “ICT’s

for future Economic and Sustainable Agricultural Systems", Perth, Australia, 2014.

[11] S. Jabjone and S. Wannasang, "Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand", International Journal of Computer and Electrical Engineering, 6(2), pp. 162-166, 2014.