CROP YEID PREDICTION USING NEURAL NETWORKS & MACHINE LEARNING

[1] CH . MALLIKARJUNARAO, [2] M. HARIKA

[1.] Professor, Department of Computer Science and Engineering, Gokaraju Rangaraju Institute Of Engineering And Technology,(TS).India.
Email-: cmrao@griet.ac.in

[2.] M.Tech Student Department of Computer Science and Engineering, Gokaraju Rangaraju Institute Of Engineering And Technology,(TS).India.
Email-: harika20241d5807@grietcollege.com

ABSTRACT

e impact of climate change in India, most of the agricultural crops are being badly affected in terms of their performance over a period of the last two decades. Predicting the crop yield in advance of its harvest would help the policy makers and farmers for taking appropriate measures for marketing and storage. This project will help the farmers to know the yield of their crop before cultivating onto the agricultural field and thus help them to make the appropriate decisions. It attempts to solve the issue by building a prototype of an interactive prediction system. Implementation of such a system with an easy-to-use web based graphic user interface and the machine learning algorithm will be carried out. The results of the prediction will be made available to the farmer. Thus, for such kind of data analytics in crop prediction, there are different techniques or algorithms, and with the help of those algorithms we can predict crop yield. Random forest algorithm is used. By analyzing all these issues and problems like weather, temperature, humidity, rainfall, moisture, there is no proper solution and technologies to overcome the situation faced by us. In India, there are many ways to increase the economic growth in the field of agriculture. Data mining is also useful for predicting crop yield production. Generally, data mining is the process of analyzing data from various viewpoint and summarizing it into important information. Random forest is the most popular and powerful supervised machine learning algorithm capable of performing both classification and regression tasks, that operate by constructing a multitude of decision trees during training time and generating output of the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

KEYWORDS: RNN,LSTM,CROP YEID

DOI Number: 10.14704/nq.2022.20.12.NQ77037 **NeuroQuantology 2022; 20(12): 529-536**

INTRODUCTION

The history of agriculture in India[1] dates back to the Indus Valley Civilization Era. India ranks second in this sector. Agriculture and allied sectors like forestry and fisheries account for 15.4 percent of the GDP (gross domestic product) with about 31 percent of the workforce. India ranks first globally with the ighest net cropped area followed by US and China. Agriculture is demographically the broadest economic sector and plays a significant role in the overall socio-economic fabric of India. Due to the revolution in industrialization, the economic contribution of agriculture to India's GDP is steadily declining with the country's broad-based economic growth.

eISSN 1303-515

India is ranked 2nd worldwide in farm output . Agriculture and allied sectors like forestry and fisheries accounted for 16.6 percent of the GDP 2009, about 50 percent of the overall workforce . The monetary contribution of agriculture to India‘s GDP is regularly declining. The crop yield of plants relies on different factors like on climatic, geographical, organic, political and financial elements. For farmers, it is difficult when there is more than one
crop to grow especially when the market prices are unknown to them. Citing the Wikipedia statistics, the farmer suicide rate in India has ranged between 1.4 and 1.8 per 100000 total population, over a 10-year period through 2005. While 2014 saw 5650 farmer suicides, the figure crossed 8000 in 2015 .

In recent times, it has become inevitable to use technology to create awareness about cultivation. The seasonal climatic conditions are

tions of crops by studying the factors such as rainfall, temperature, area (in hectares), season, etc. The system also helps in suggesting whether a particular time is the right one to use fertilizers. Crop yield prediction is an important agricultural problem. Every farmer always tries to know how much yield will be produced and whether it meets their expectations. In the past, yield prediction was calculated by analyzing a farmer's previous experience on a particular crop. The

Agricultural yield is primarily dependent on weather conditions pests and planning of harvest operation. Accurate information about the history of crop yield is an important thing for making decisions related to agricultural risk management.

II LITERATURE SURVEY

2.1Predicting yield of the crop using machine learning algorithm

also being changed against the fundamentaassets like soil, water and air which lead to insecurity of food. In a scenario, crop yield rate is falling short of meeting the demand consistently and there is a need for a smart system which can

solve the problem of decreasing crop yield. Therefore, to eliminate this problem, we propose a system which will provide crop selection based on economic and environmental factors to reap the maximum yield out of it for the farmers which will sequentially help meet the elevating demands for the food supplies in the country. The proposed system uses machine learning to make the predictions. The system will provide crop yield and crop selection based on weather attributes suitable for the crop to get the maximum yield out of it for the farmers. The systemmakes predictions of the pro

AUTHORS: P.Priya, U.Muthaiah&M.Balamurugan

The agriculture plays a dominant role in the growth of the country’s economy.Climate and other environmental changes has become a major threat in the agriculture field. Machine learning (ML) is an essential approach for achieving practical and effective solutions for this problem. Crop Yield Prediction involves predicting yield of the crop from available historical available data like weather parameter,soil parameter and historic crop yield.This paper focus on predicting the yield of the crop based on the existing data by using Random Forest algorithm. Real data of Tami lnadu were used for building the models and the models were tested with samples.The prediction will helps to the farmer to predict the yield of the crop before cultivating onto the agriculture field. To predict the crop yield in future accurately

eISSN 1303-515

Random Forest, a most powerful and popular supervised machine learning algorithm is used.

2.2 Applications of machine learning techniques in agricultural crop production: a review

AUTHORS: Mishra .s, Mishra .D and Santra .G. H

This paper has been prepared as an effort to reassess the research studies on the relevance of machine learning techniques in the domain of agricultural crop production. Methods/Statistical Analysis: This method is a new approach for production of agricultural crop management. Accurate and timely forecasts of crop production are necessary for important policy decisions like import-export, pricing marketing distribution etc. which are issued by the directorate of economics and statistics. However one has understand that these prior estimates are not the objective estimates as these estimate requires lots of descriptive assessment based on many different qualitative factors. Hence there is a requirement to develop statistically sound objective prediction of crop production. That development in computing and information storage has provided large amount of data. Findings: The problem has been to intricate knowledge from this raw data, this has lead to the development of new approach and techniques such as machine learning that can be used to unite the knowledge of the data with crop yield evaluation. This research has been intended to evaluate these innovative techniques such that significant relationship can be found by their applications to the various variables present in the data base. Application/Improvement: The few techniques like artificial neural networks, Information Fuzzy Network, Decision Tree, Regression Analysis, Bayesian belief network. Time series analysis, Markov chain model, k-means clustering, k nearest neighbor, and support vector machine are applied in the domain of agriculture were presented.

2.3 A Model for Prediction of Crop Yield.

AUTHORS: Manjula.E

Data Mining is emerging research field in crop yield analysis. Yield prediction is a very important issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data. Data mining techniques are the better choice for this purpose. Different Data Mining techniques are used and evaluated in agriculture for estimating the future year's crop production. This research proposes and implements a system to predict crop yield from previous data. This is achieved by applying association rule mining on agriculture data. This research focuses on creation of a prediction model which may be used to future prediction of crop yield. This paper presents a brief analysis of crop yield prediction using data mining technique based on association rules for the selected region i.e. district of Tamil Nadu in India. The experimental results shows that the proposed work efficiently predict the crop yield production.

2.4 Agricultural crop yield prediction using artificial neural network approach

AUTHORS: Dahikar, S. S, Rode and S. V.

By considering various situations of climatologically phenomena affecting local weather conditions in various parts of the world. These weather conditions have a direct effect on crop yield. Various researches have been done exploring the connections between

eISSN 1303-515

large-scale climatologically phenomena and crop yield. Artificial neural networks have been demonstrated to be powerful tools for modeling and prediction, to increase their effectiveness. Crop prediction methodology is used to predict the suitable crop by sensing various parameter of soil and also parameter related to atmosphere. Parameters like type of soil, PH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulphur, manganese, copper, iron, depth, temperature, rainfall, humidity. For that purpose we are used artificial neural network (ANN).

2.5 Predictive ability of machine learning methods for massive crop yield prediction.

AUTHORS: GonzlezSnchez. A, Frausto Sols.J and Ojeda Bustamante. W

An important issue for agricultural planning purposes is the accurate yield estimation for the numerous crops involved in the planning. Machine learning (ML) is an essential approach for achieving practical and effective solutions for this problem. Many comparisons of ML methods for yield prediction have been made, seeking for the most accurate technique. Generally, the number of evaluated crops and techniques is too low and does not provide enough information for agricultural planning purposes. This paper compares the predictive accuracy of ML and linear regression techniques for crop yield prediction in ten crop datasets. Multiple linear regression, M5-Prime regression trees, perceptron multilayer neural networks, support vector regression and k-nearest neighbor methods were ranked. Four accuracy metrics were used to validate the models: the root mean square error (RMS), root relative square error (RRSE), normalized mean absolute error (MAE), and correlation factor (R). Real data of an irrigation zone of Mexico were used for building the models. Models were tested with samples of two consecutive years. The results show that M5- Prime and k-nearest neighbor techniques obtain the lowest average RMSE errors (5.14 and 4.91), the lowest RRSE errors (79.46% and 79.78%), the lowest average MAE errors (18.12% and 19.42%), and the highest average correlation factors (0.41 and 0.42). Since M5-Prime achieves the largest number of crop yield models with the lowest errors, it is a very suitable tool for massive crop yield prediction in agricultural planning.

III.EXISTING SYSTEM

The problem that the Indian Agriculture sector is facing is the integration of technology to bring the desired outputs. With the advent of new technologies and overuse of non-renewable energy resources patterns of rainfall and temperature are disturbed.

IV PROPOSED SYSTEM:

The inconsistent trends developed from the side effects of global warming make it cumbersome for the farmers to clearly predict the temperature and rainfall patterns thus affecting their crop yield productivity. In order to perform accurate prediction and handle inconsistent trends in temperature and rainfall various machine learning algorithms like RNN, LSTM, etc can be applied to get a pattern. It will complement the agricultural growth in India and all together augment the ease of living for farmers. In past, many researchers have applied machine learning techniques to enhance agricultural growth of the country.

IV.SYSTEM ARCHITECTURE

eISSN 1303-515

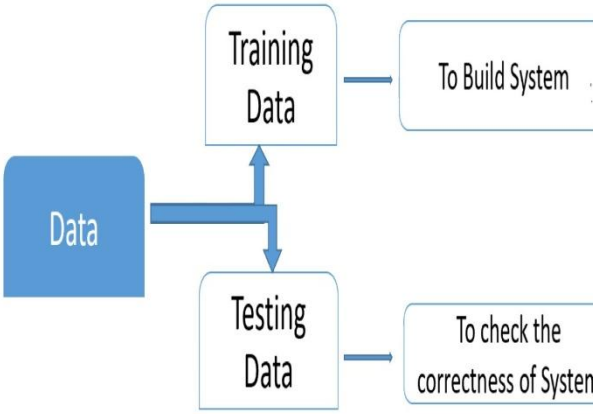
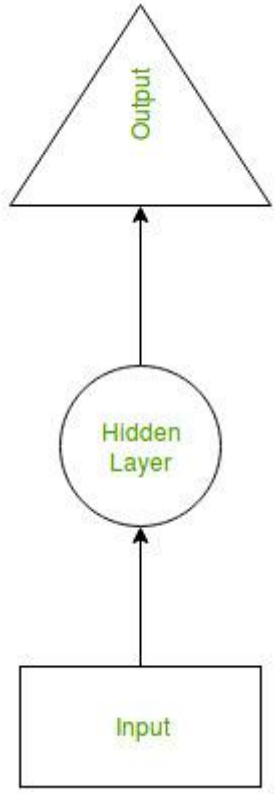


Fig1: Architecture of face expression recognition system.

ALGORITHMS:

Recurrent Neural Network(RNN) are a type of Neural Network where the **output from previous step are fed as input to the current step**. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is **Hidden state**, which remembers some information about a sequence.



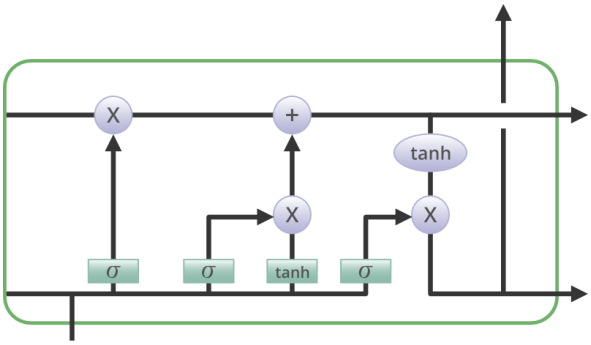
RNN have a **"memory"** which remembers all information about what has been calculated. It uses the same parameters for each input as it performs the same task on all the inputs or hidden layers to produce the output. This reduces the complexity of parameters, unlike other neural networks.

Long Short Term Memory(LSTM Algorithm) is a kind of recurrent neural network. In RNN output from the last step is fed as input in the current step. LSTM was designed by Hochreiter&Schmidhuber. It tackled the problem of long-term dependencies of RNN in which the RNN cannot predict the word stored in the long-term memory but can give more accurate predictions from the recent information. As the gap length increases RNN does not give an efficient performance. LSTM can by default retain the information for a long period of time. It is used for processing,

eISSN 1303-515

predicting, and classifying on the basis of time-series data.

Structure of LSTM:
LSTM has a chain structure that contains four neural networks and different memory blocks called cells.



MODULES:

Upload Crop Dataset
The crop production dataset that is used to predict the name and yield of the crop is fed into classification and regression algorithms.

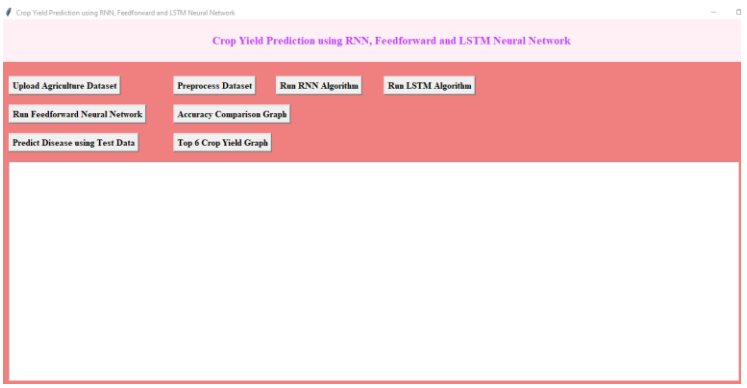
Preprocess Dataset
Experiments were conducted on Indian government dataset and it has been established that Random Forest Regressor gives the highest yield prediction accuracy. Sequential model that is Simple Recurrent Neural Network performs better on rainfall prediction while LSTM is good for temperature prediction. By combining rainfall, temperature along with other parameters like season and area, yield prediction for a certain district can be made.

Train Machine Learning
This focuses on district wise yield prediction according to the crop sown in the district. Yield is being predicted for given crops district wise and crops with best yield.

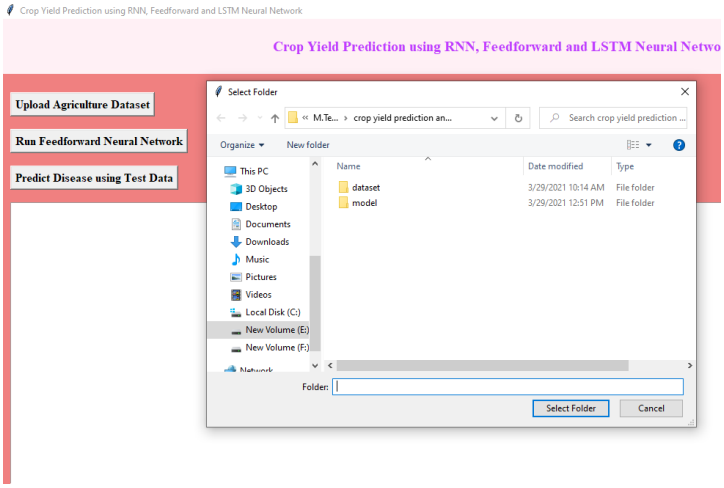
Upload Test Data &Predict Yield
Results reveal that Random Forest is the best classifier when all parameters are combined. This will not only help farmers in choosing the right crop to grow in the next season but also bridge the gap between technology and the agriculture sector.

:

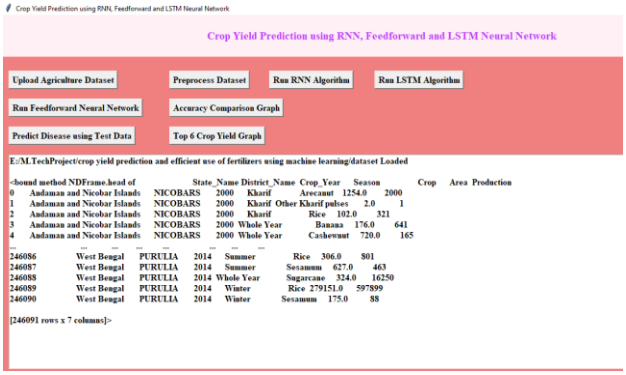
To run project double click on 'run.bat' file to get below screen



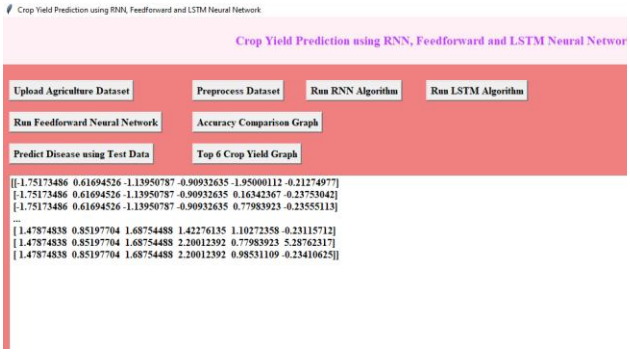
In above screen click on 'Upload Crop Dataset' button to upload dataset



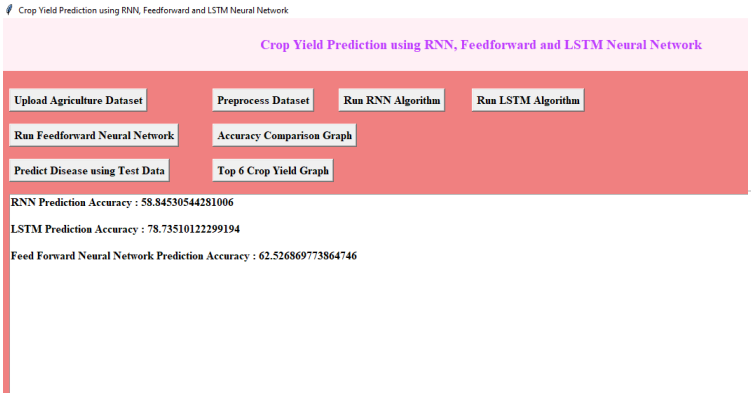
In above screen selecting and uploading 'Dataset.csv' file and then click on 'Open' button to load dataset and to get below screen



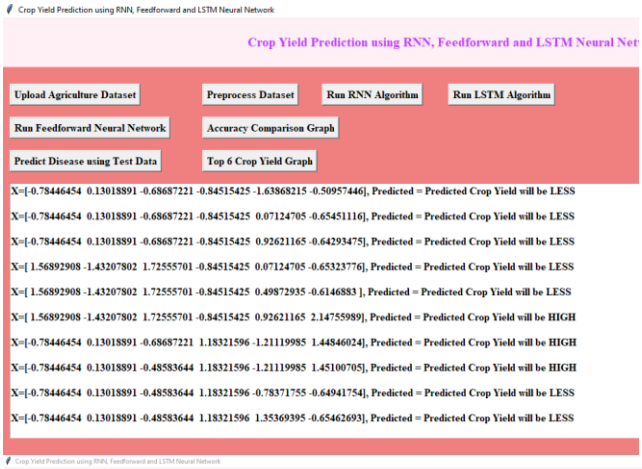
In above screen dataset loaded and we can see dataset contains some non-numeric values and ML will not take non-numeric values so we need to preprocess dataset to convert non-numeric values to numeric values by assigning ID to each non-numeric value. So click on 'Preprocess Dataset' button to process dataset



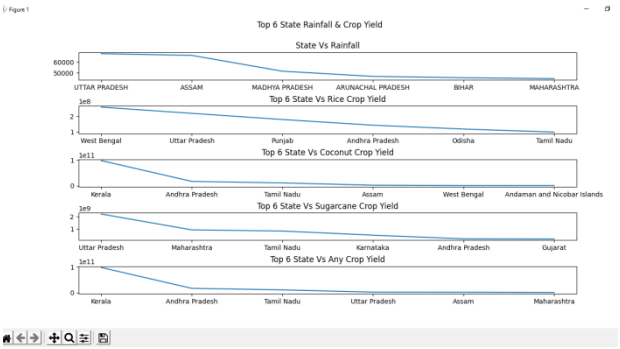
In above screen all non-numeric values converted to numeric format and in below lines we can see dataset contains total 246091 records and application using (80%) 196872 records to train ML and using (20%) 49219 records to test ML prediction error rate (RMSE (root mean square error)). Now click on 'Train Machine Learning Algorithm' button to train Decision Tree Machine learning algorithm on above dataset and then calculate prediction error rate



In above screen ML is trained and we got prediction error rate as 0.067% and now Decision Tree model is ready and now click on 'Upload Test Data & Predict Yield' button to upload test data and then application will predict production



In above screen selecting and uploading 'test.csv' file and then click on 'Open' button to load test data and then application will give below prediction result



In above screen each test record is separated with newline and in above screen in square bracket we can see test data values and after square bracket we can see predicted production and after that we can see predicted YIELD per acre. So each test record and its prediction is separated with newline.

VII. CONCLUSION

The paper presented the various machine learning algorithms for predicting the yield of the crop on the basis of temperature, rainfall, season and area. Experiments were conducted on Indian government dataset and it has been established that Random Forest Regressor gives the highest yield prediction accuracy. Sequential model that is Simple Recurrent Neural Network performs better on rainfall prediction while LSTM is good for temperature prediction. By combining rainfall, temperature along with other parameters like season and area, yield prediction for a certain district can be made. Results reveal that Random Forest is the best classifier when all parameters are combined. This will not only help farmers in choosing the right crop to grow in the next season but also bridge the gap between technology and the agriculture sector.

The future work is focused on providing the sequence of crops to be grown depending on the soil and weather conditions and to update the datasets time to time to produce accurate predictions. The Future Work targets a fully automated system that will do the same. Another functionality that we are trying to implement is to provide the correct fertiliser for the given crop and location. To implement this through study of fertilisers and their relationship with soil and climate is required..

IX. REFERENCES

1. Agriculture Role on Indian Economy Madhusudhan L - https://www.omicsonline.org/open-access/agriculture-role-on-indianeconomy-2151-6219-1000176.php?aid=62176

2. Priya, P., Muthaiah, U., Balamurugan, M. International Journal of Engineering Sciences Research Technology Predicting Yield of the Crop Using Machine Learning Algorithm.

4.Mishra, S., Mishra, D., Santra, G. H. (2016). Applications of machine learning techniques in agricultural crop production: a review paper.Indian J. Sci. Technol,9(38), 1-14.

4. Manjula, E., Djodiltachoumy, S. (2017). A Model for Prediction of Crop Yield.International Journal of Computational Intelligence and Informatics,6(4), 2349-6364.

5. Dahikar, S. S., Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach.International journal of innovative research in electrical, electronics, instrumentation and control engineering,2(1), 683-686.

6. GonzlezSnchez, A., Frausto Sols, J., Ojeda Bustamante, W. (2014). Predictive ability of machine learning methods for massive crop yield prediction.

7.Mandic, D. P., Chambers, J. (2001). Recurrent neural networks for prediction: learning algorithms, architectures and stability. John Wiley Sons, Inc..

8. Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory.Neural computation, 9(8), 1735-1780.

9.Sak, H., Senior, A., Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling.In Fifteenth annual conference of the international speech communication association.

eISSN 1303-515