CROP YIELD PREDICTION UTILIZING MULTIMODAL DEEP LEARNING

1V.Mallesi M.Tech., Head of the Department, Department of CSE, Bheema Institute Of Technology And Science - [BITS], csehod.bitsadoni@gmail.com

2V Kusuma Sri Btech, Department of CSE, Bheema Institute Of Technology And Science - [BITS], kusumasrivitikipati4@gmail.com

ABSTRACT: Agriculture is one of the major and the least paid occupation in India. Machine learning can bring a boom in the agriculture field by changing the income scenario through growing the optimum crop. This paper focuses on predicting the yield of the crop by applying various machine learning techniques. The outcome of these techniques is compared on the basis of mean absolute error. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get the maximum yield by considering factors like temperature, rainfall, area, etc.

Keywords *– Crop yield prediction, deep learning*

1. 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 highest 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. 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. 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. This paper focuses on predicting the yield of the crop by applying various machine learning techniques. The outcome of these techniques is compared on the basis of mean absolute error. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get

the maximum yield by considering factors like temperature, rainfall, area, etc.

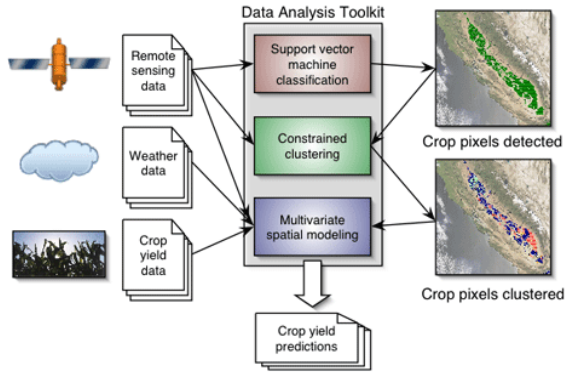

Fig.1: Example figure

Many studies have used machine learning techniques such as regression tree, random forest, multivariate regression, association rule mining, and artificial neural networks for crop yield prediction. Machine learning models treat the output, crop yield, as an implicit function of the input variables such as weather components and soil conditions, which could be a very complex and nonlinear function. Jeong et al. applied random forest and multiple linear regression for yield prediction of wheat, maize, and potato. They found that random forest was highly capable of predicting crop yields and outperformed multiple linear regression. Fukuda et al. also used random forest for predicting mango fruit yields in response to water supply under different irrigation regimes, and found that random forest was applicable for mango yield prediction with a specific focus on water management. Liu et al. applied artificial neural networks to approximate a nonlinear function to relate the corn yield to input variables such as weather, soil, and management practices. Ransom et al. evaluated machine learning methods for corn nitrogen recommendation tools using soil and weather information. Drummond et al. investigated stepwise multiple linear regression, projection pursuit regression, and artificial neural networks to predict the grain yield based on the soil properties. Shahhosseini et al. predicted corn yield and nitrate loss using machine learning algorithms such as random forest and multiple linear regression. Awad designed a mathematical optimization model to predict potato yield using the biomass calculated by the model. Jiang et al. applied artificial neural network and multiple linear regression for estimating winter wheat yields based on the remotely sensed and climate data, and found that artificial neural network model outperformed the multiple linear regression. Prasad et al. used piecewise linear regression method with breakpoint to predict corn and soybean yields based on remote sensing data and other surface parameters. Romero et al. applied several machine learning methods such as decision tree and association rule mining for the classification of yield components of durum wheat and showed that association rule mining method obtained the best performance across all locations.

2. LITERATURE REVIEW

PREDICTING YIELD OF THE CROP USING MACHINE LEARNING ALGORITHM

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

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

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.

A Model for Prediction of Crop Yield.

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.

Agricultural crop yield prediction using artificial neural network approach

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 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).

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

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.

3. METHODOLOGY

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. 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. 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.

This paper focuses on the practical application of machine learning algorithms and its quantification. The work presented here also takes into account the inconsistent data from rainfall and temperature datasets to get a consistent trend. Crop yield prediction is determined by considering all the features in contrast with the usual trend of determining the prediction considering one feature at a time.

Advantages:

∟ Achieving the maximum crop at minimum yield is the ultimate Aim of the project. Early detection of problems and management of that problems can help the farmers for better crop yield. For the better understanding of the crop yield, we need to study of the huge data with the help of machine learning algorithm so it will give the accurate yield for that crop and suggest the farmer for a better crop.

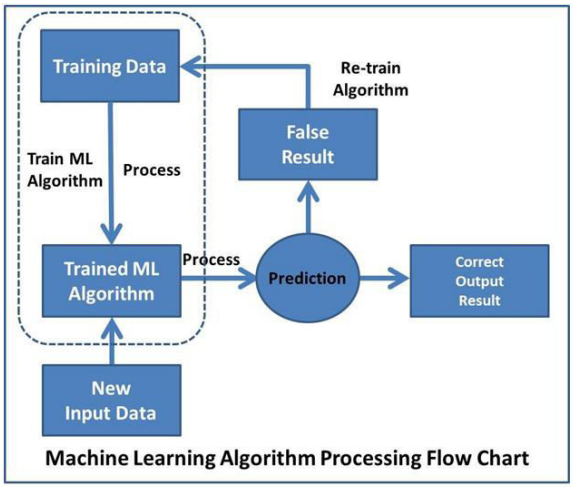

Fig.2: System architecture

MODULES DESCRIPTION:

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.

4. ALGORITHM

Long short-term memory (LSTM):

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single

data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic or IDSs (intrusion detection systems).

A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three *gates* regulate the flow of information into and out of the cell.

LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications

Training:

An RNN using LSTM units can be trained in a supervised fashion, on a set of training sequences, using an optimization algorithm, like gradient descent, combined with backpropagation through time to compute the gradients needed during the optimization process, in order to change each weight of the LSTM network in proportion to the derivative of the error (at the output layer of the LSTM network) with respect to corresponding weight.

A problem with using gradient descent for standard RNNs is that error gradients vanish exponentially quickly with the size of the time lag between important events. However, with LSTM units, when error values are back-propagated from the output layer, the error remains in the LSTM unit's cell. This "error carousel" continuously feeds error back to each of the LSTM unit's gates, until they learn to cut off the value.

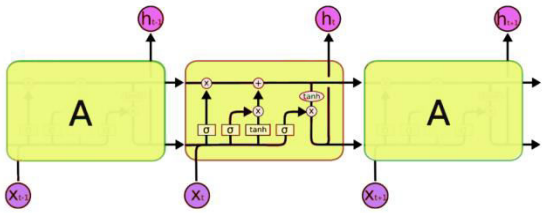

Fig.3: LSTM model

5. EXPERIMENTAL RESULTS

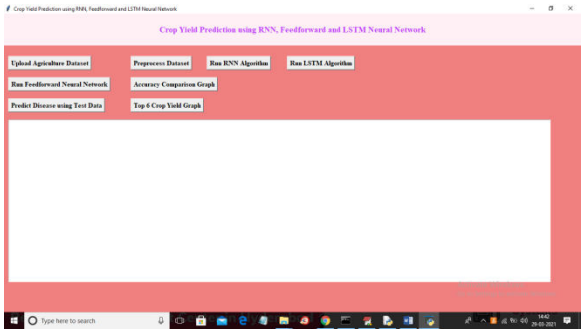

Fig.4: Home screen

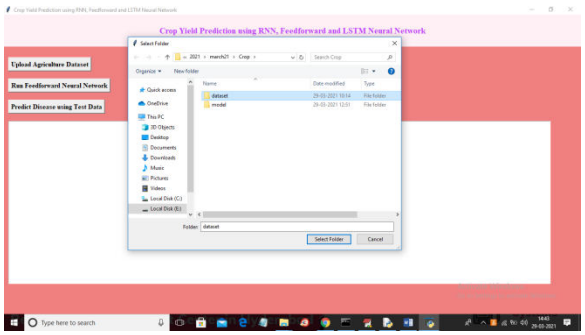

Fig.5: Upload dataset

Fig.6: Dataset loaded

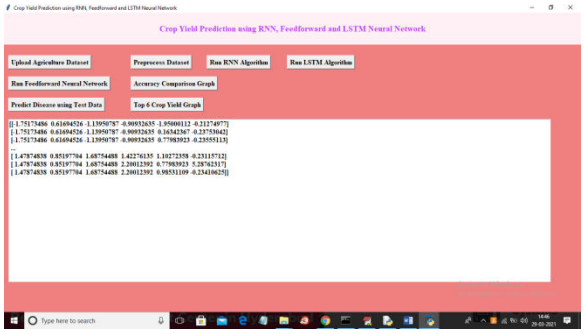

Fig.7: Preprocess dataset

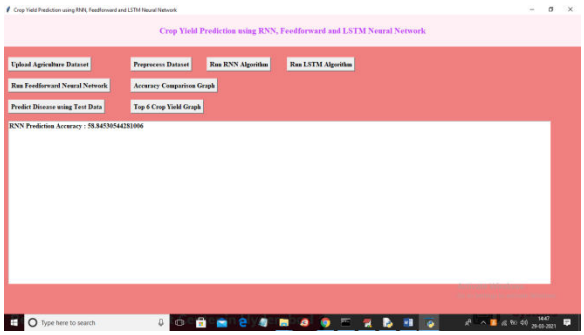

Fig.8: RNN algorithm

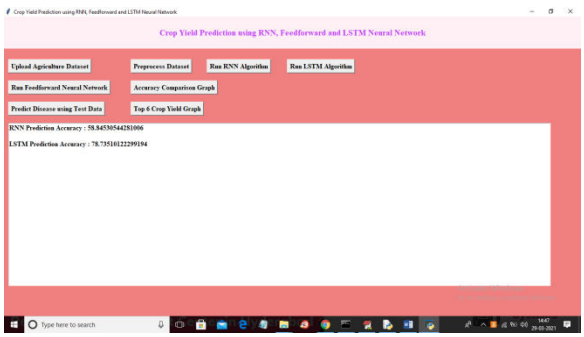

Fig.9: LSTM algorithm

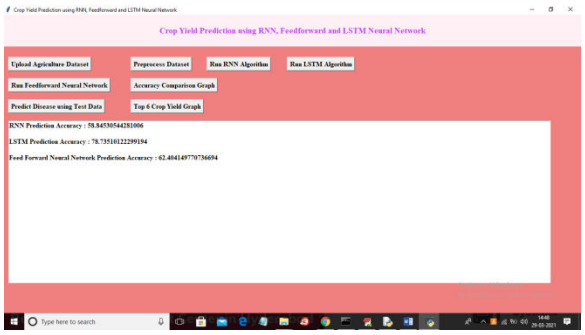

Fig.10: Run feedforward neural network

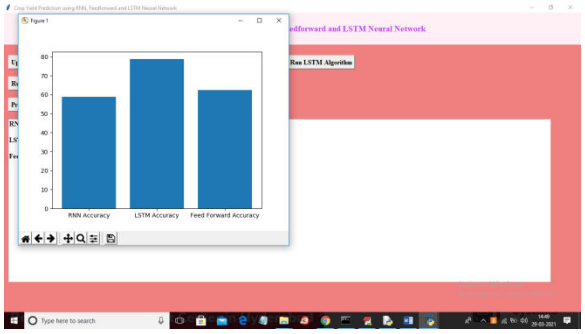

Fig.11: Accuracy comparison graph

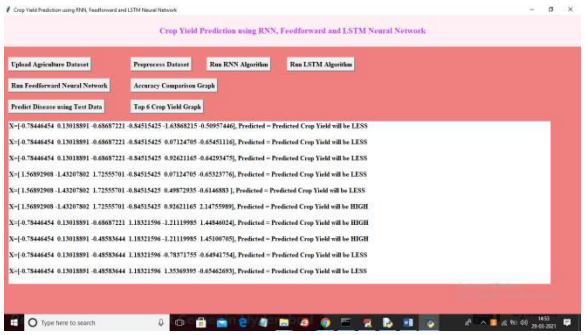

Fig.12: Predict disease using test data

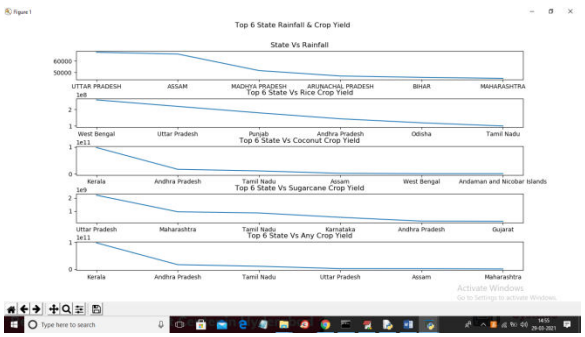

Fig.13: Top 6 crop yield graph

6. CONCLUSION

This paper focuses on predicting the yield of the crop by applying various machine learning techniques. The outcome of these techniques is compared on the basis of mean absolute error. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get the maximum yield by considering factors like temperature, rainfall, area, etc.

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.

3.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), 2349-6363.

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. Gonzlez Snchez, 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.

10. Liaw, A., Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.