CROP GUIDANCE USING MACHINE LEARNING

BOPPA ABHINAV
Department of CSE
boppaabhinav@gmail.com
Teegala Krishna Reddy Engineering College, Hyderabad

T.MAHESH
Department of CSE
thatikondamahesh950@gmail.com
Teegala Krishna Reddy Engineering College, Hyderabad

M.PRANAY
Department of CSE
pranaykumar4433@gmail.com
Teegala Krishna Reddy Engineering College,Hyderabad

J.RAGHAVENDER
Department of CSE
jraghavendar98@gmail.com
Teegala Krishna Reddy Engineering College,Hyderabad

B.NARSINGAM
Assistant Professor
Department of CSE
Teegala Krishna Reddy Engineering College,Hyderabad

Abstract - Agriculture is the field which plays a very important role in improving our country's economy. Agriculture is that the one which gave birth to civilization. India is an agricultural country and its economy is largely based upon crop productivity. Hence, we are able to say that agriculture will be the backbone of all business in our country. Selecting every crop is extremely important in agricultural planning. The choice of crops will rely on the various parameters like market value, production rate and also the different government policies. Many changes are required within the agriculture field to boost changes in our Indian economy. We will improve agriculture by using machine learning techniques which are applied easily within the farming sector. Together with all advances within the machines and technologies employed in farming, useful and accurate information about different matters also plays a big role in it. The concept of this project is to implement the crop selection method so this method helps in solving many agriculture and farmers problems. This improves our Indian economy by maximizing the yield rate of crop production.

Index Terms - Crop yield prediction, Machine learning techniques, Random Forest, Decision Tree, Supervised Learning

I. INTRODUCTION

The impact of temperature 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 before its harvest would help the policy makers and farmers for taking appropriate measures for marketing and storage. This project will help the farmers to understand the yield of their crop before cultivating onto the agricultural field and this will help to suitable findings. It attempts to unravel the problem by building a prototype of an interactive prediction system. Implementation of such a system with an easy-to-use web based graphic program and therefore the machine learning algorithms are administrated. The results of the predictions are going to be made available to the farmer. Thus, for such styles of data analytics in crop prediction, there are different techniques or algorithms, and with the assistance of these algorithms we are able to predict crop yield. By analyzing these issues, and problems like weather, temperature, humidity, rainfall, moisture, there's no proper solution and technologies to beat true faced by us. In India, there are some ways to extend the economic process within the field of agriculture. Data preprocessing is additionally useful for predicting crop yield production. Generally, data processing is that the process of analyzing data from various viewpoints and summarizing it into important information. Random forest is that the most popular and powerful supervised machine learning algorithm capable of performing both classification and regression tasks, that operate by constructing a mess of decision trees during training time and generating output of the category that's the mode of the classes(classification) or mean prediction (regression) of the individual trees.

Scope of the Project

The scope of the project is to determine the crop yield of an area by considering dataset with some features which are important or related to crop production such as temperature, moisture, rainfall, and production of the crop in previous years. To predict a continuous value, regression models are used. It is a supervised technique. The coefficients are preprocessed and fit into the trained data during training and construction the regression model. The main focus here is to reduce the cost function by finding the best fit-line. The output function facilitates in error measurement. During training period, error between the predicted and actual values is reduced in order to minimize error function. Python is used for this project by pycharm platform for stimulation.

II. LITERATURE SURVEY

Machine learning methods for crop yield prediction and climate change impact assessment in agriculture

Crop yields are highly reliant on the weather. A increasing body of empirical work attempts to simulate this connection in order to forecast the effects of climate change on the industry. We provide a yield modelling technique that employs a semi-parametric version of a deep neural network to account for complicated nonlinear connections in high-dimensional datasets, as well as known parametric structure and undiscovered cross-sectional variability. We demonstrate that this technique is applicable for both traditional statistical approaches and fully-nonparametric neural networks in forecasting yields of years withheld during model training using data on maize

yield from the US Midwest. We show large negative impacts of temperature change on corn yield using scenarios from a set of climate models, but they are less severe than impacts projected using classical statistical methods. In particular, our approach is less pessimistic within the warmest regions and also the warmest scenarios.

Smart Farming Crop Yield Prediction using Machine Learning

For over two decades, India's agriculture industry has been in decline. Suicide incidents have been on the rise in recent years. This stems from a lack of adequate understanding as well as the routine farming practices used by farmers. Crop production is influenced by a variety of seasonal, economic, and biological trends. Changes in these patterns that are catastrophic might cause farmers to lose a lot of money. Smart agricultural techniques, i.e. putting technology into day-to-day farming, will mitigate these dangers. The major goal of the study is to get helpful insights on agricultural yield prediction, prognostication, crop type planting, and crop cost forecasting. Experimental analysis is carried out on the statistical agriculture dataset. The data is preprocessed and divided into two categories: training and testing. Then, for improved classification results, appropriate classification algorithms such as Support Vector Machine (SVM) and Random Forest are utilised.

A Model for Prediction of Crop Yield

Data Mining is an emerging research field in crop yield analysis. Yield prediction may be an important issue in agriculture. Any farmer is fascinated by knowing what quantity yield he's near to expect. Within the past, yield prediction was performed by considering farmer's experience on a specific field and crop. The yield prediction may be a major issue that is still to be solved supported available data. Data processing techniques are the higher choice for this purpose. Different data processing techniques are used and evaluated in agriculture for estimating the longer-term year's crop production. This research proposes and implements a system to predict crop yield from previous data. This is often achieved by applying association rule mining on agriculture data. This research focuses on creation of a prediction model which can be used for future prediction of crop yield.

III Methodology

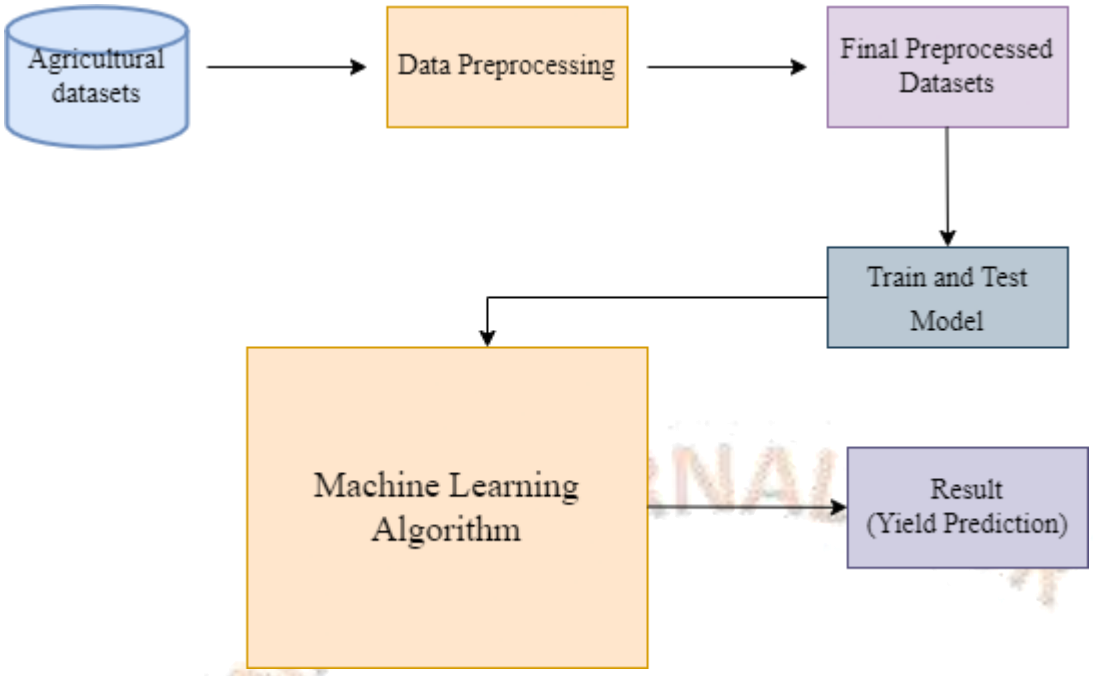
Machine learning mainly consists of three learning methods, namely supervised learning, reinforcement learning and unsupervised learning methods of training a model. Supervised learning is a learning method which maps known input resulted into output which maps from input to output. But in case of unsupervised leaning we would not know targeted output in this learning we should train the model in order to get desired output.

Existing System

- Deep Neural Networks used to predict yield, check yield, and yield difference of corn hybrids from genotype and environment data.
- The training data included three sets, crop genotype, yield performance, and environment.
- Complex model structures make it hard to produce testable hypotheses that could potentially provide biological insights.

Proposed System

- Our proposed system is a mobile application which predicts name of the crop as well as calculate its corresponding yield.
- Name of the crop is determined by several features like temperature, humidity, wind-speed, rainfall etc. and yield is determined by the area and production.
- Providing the user to select based on which field he/she wants to perform the analysis.
- Analyzing the data present with us based on the provided user requirement.



System Architecture

IV Algorithms

Decision Tree:
It's a supervised learning method that's mostly used to solve classification issues. It works for both categorical and continuous dependent variables, which is surprising. We divide the population into two or more homogenous sets using this technique. This is based on the most important characteristics/independence.

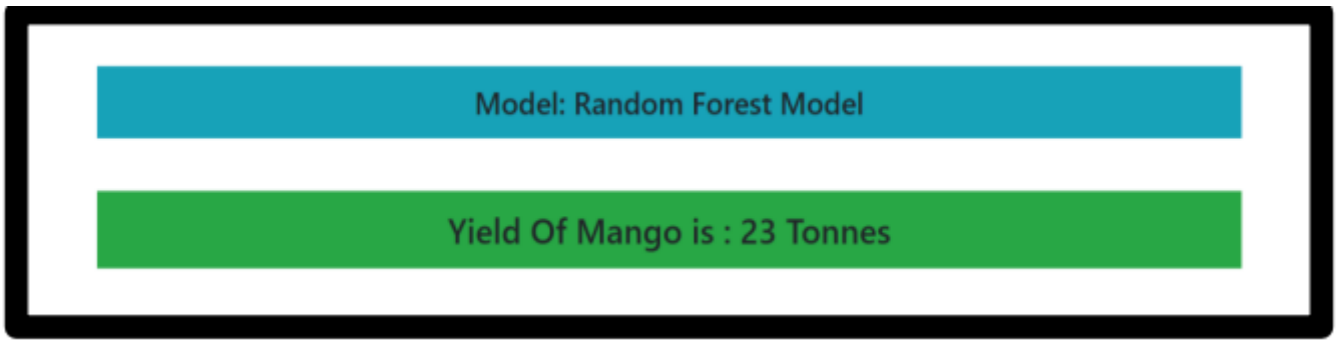
Random Forest
An ensemble of decision trees is referred to as a Random Forest. We have a collection of decision trees (referred to as "Forest") in Random Forest. Each tree offers a categorization to a new object based on characteristics, and we say the tree "votes" for that class. The categorization with the highest votes is chosen by the forest (over all the trees in the forest). In Each tree is planted & grown as follows:
➢ If the number of cases in the training set is N, then sample of N cases is taken at random but with replacement. This sample is going to be the training set for growing the tree.
➢ If there are M input variables, a number m<<M is specified such that at each node, m variables are selected at random out of the M and the best split on these m is employed to split the node. M value is constant at the time of forest growing.

V RESULTS

The Recommended crops are below:

Blackgram	20.0 %
Maize	5.0 %
Papaya	15.0 %
Pigeonpeas	60.0 %

PREDICTED CROPS



.PREDICTED YIELD

III. CONCLUSIONS

The suggested method considers soil N, P, K, humidity, temperature, pH, and rainfall variables to predict which crops will be the most productive in those conditions. Because the system uses data mining techniques to list all of the available crops, it assists the farmer in deciding which crop to produce in their location. This approach therefore assists the farmer in determining the most lucrative crop as well as identifying new crops that may be planted that the farmer has not before farmed. The created homepage is user-friendly, and prediction accuracy is greater than 75% in all of the crops and districts studied, suggesting increased forecast accuracy. Any user of their choice of crop can utilise the user-friendly web page designed for forecasting crop production by supplying meteorological data for that location.

IV. REFERENCES

1. Agriculture Role on Indian Economy Madhusudhan L https://www.omicsonline.org/openaccess/ 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), 114.
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.
11. Chen, T., Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acmsigkdd international conference on knowledge discovery and data mining (pp.785-794). ACM. 38
12. Cover, T. M., Hart, P. E. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27. 13. Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., Klein, M. (2002). Logistic regression. New York: Springer-Verlag.
14. Seber, G. A., Lee, A. J. (2012). Linear regression analysis (Vol. 329). John Wiley Sons.
15. urada, J. M. (1992). Introduction to artificial neural systems (Vol. 8). St. Paul: West publishing company.
16. http://www.indiawaterportal.org/met data/
17. District-wise, season-wise crop production statistics Vkhullar - https://data.gov.in/catalog/district-wise-season-wise-crop-productionstatistic