A Review on Crop Yield Prediction Based on Indian Agriculture Using Machine Learning

Nitesh patel[1], Prof. Sheela singh[2]
[1, 2] Dept of Computer Science
[1, 2] Saraswati Higher Education and Technical College of Engineering Babatpur, Varanasi Uttar Pradesh221106.

Abstract- *In India, we all know that Agriculture is the backbone of the country. This paper predicts the yield of almost all kinds of crops that are planted in India. This script makes novel by the usage of simple parameters like State, district, season, area and the user can predict the yield of the crop in which year he or she wants to. The paper uses advanced regression techniques like Kernel Ridge, Lasso and ENet algorithms to predict the yield and uses the concept of Stacking Regression for enhancing the algorithms to give a better prediction. The work proposes to help farmers check the soil quality depending on the analysis done based on data mining approach. Thus the system focuses on checking the soil quality to predict the crop suitable for cultivation according to their soil type and maximize the crop yield with recommending appropriate fertilizer.*

Keywords- Crop yield prediction, Lasso, Kernel Ridge, ENet, Stacked Regression.

I. INTRODUCTION

Agriculture is a significant subject of interest to society since it produces a huge part of food. Many nations are still experiencing hunger as a result of a lack of food and a rising population. Expanding food production is a convincing method of eradicating starvation. Developing food security and reducing hunger by 2030 are major goals for the United Nations. Crop protection, land evaluation, and crop yield forecast are therefore increasingly important to world food production. A country's policymakers rely on accurate forecasts to make suitable export and import evaluations in order to strengthen national food security. Dongxiao Yu was the assistant editor who coordinated the evaluation of this article and approved it for publication. Farmers may also use yield forecasting to make financial and managerial choices. Agricultural supervision, particularly crop production monitoring, is critical in determining a region's food security. Crop production forecasting, on the other hand, is very difficult due to a variety of complicated factors. Crop production is primarily determined by meteorological circumstances, soil quality, landscapes, insect infestations, water quality and availability, genotype, harvest activity planning, and so on.

Crop yield processes and strategies change over time and are deeply non-linear in nature and complicated owing to the integration of a large number of associated components, which are defined and influenced by non-arbitrate runs and external causes.

Typically, a significant portion of the agricultural framework cannot be demarcated in a basic stepwise computation, particularly when dealing with complicated, incomplete, confusing, and contentious statistics. Many research now show that machine learning algorithms have a higher potential than traditional statistics. Machine learning is a branch of artificial intelligence in which computers may be educated without explicit programming. With outstanding predicting power, these algorithms resolve non-linear or linear agricultural contexts. Machine learning agriculture frameworks derive their strategies from the learning process. These methods need overtraining to complete a certain goal. Following the training phase, the model makes assumptions to test the information.

PROBLEM STATEMENT:

Crop yield forecasting is essential for planning and making policy decisions. Many countries still depend on ground-based inspections and reports to monitor crops and anticipate output. These processes are capricious, costly, and time-consuming. Some of the most common concerns with current crop production projection techniques are as follows.

- The most significant issue with current crop production forecast methods is their lack of accuracy and time consumption.
- Current time series agricultural production forecast methods do not account for fluctuations caused by cycles and seasonal impacts.
- Needs extensive information to develop and test the model and also available information in agriculture is sparse and incomplete in existing simulation model.
- Only a few research on agricultural production prediction utilizing the present decision tree approach have been conducted.

- The prediction error value is another major issue in crop production forecast or estimating systems.
- These are the key disadvantages of present studies that urge us to conduct crop yield prediction research.

Aim

The suggested approach considers soil, weather, and previous year production Aim and advises which are the most lucrative crops that may be planted under the provided environmental conditions. Because the system includes all conceivable crops, it assists the farmer in determining which crop to produce. Furthermore, this system considers previous data production, which will assist the farmer in gaining insight into the demand and pricing of different crops in the market. Because this approach will cover the most sorts of crops, farmers may learn about crops that they have never grown before.

Objective

- To increase the accuracy of crop yield prediction as well as provide an easy to use User Interface.
- To analyse different climatic parameters (cloud cover, rainfall, temperature).
- To derive simple decision models to demonstrate the usefulness of the stochastic yield projections in meeting specified agricultural policy goals.
- To obtain estimates of aggregate physical production functions for the yields of various crops in specified states, considering various technological factors and a newly developed meteorological weather index as inputs.

Scope of study:

The project's aim is to calculate an area's agricultural output by taking into account datasets containing certain parameters that are relevant or connected to crop production, such as temperature, moisture, rainfall, and crop productivity in past years. Regression models are used to forecast a continuous value. It is a controlled procedure. It focuses on the development of a prediction model that may be used to forecast agricultural yield output.

When it comes to crop output, most of us are concerned with how it will seem at harvest. Crop yield prediction is a significant agricultural issue. Every farmer wants to know how much output he may anticipate based on his expectations. Previously, yield prediction was computed by studying a farmer's prior experience with a certain crop. Agricultural production is mostly determined by weather conditions, pests, and harvest process planning. Accurate information regarding crop production history is critical for making agricultural risk management choices.

RELATED WORK

Forecasting agriculture product plays a significant role in agriculture planning. It helps in making product storage, business strategy and risk management. There are two methods to forecast agriculture product in advance. First is statistics method such as Autoregressive Integrate Moving Average (ARIMA) and Holt-Winter and second is machine learning method such as Support vector machine and artificial neural network. These methods are comparatively study over Thailand's pacific white shrimp export data and Thailand's pro- ducked chicken data using support vector machine and ARIMA model. Where support vector method gives more accurate result than ARIMA. Moreover, machine learning methods are convenient to implement and comparably faster than statics methods. Indian agriculture is highly dependent on summer rainfall The correlation between summer rainfall and agriculture product production is studied in. This paper presents an analysis of crop-climate relationship using past crops data. Correlation analysis tells that the monsoon rainfall, Pacific and Indian Ocean sea-surface temperatures and Darwin sea- level pressure directly influence the crop production in India. Result shows that the state-level crop production statistics and sub divisional monsoon rainfall are consistent with the all-India result, except few cases. Moreover, the impact of sub divisional monsoon rainfall related to El Niosouthern oscillation and the Indian Ocean sea-surface temperatures have seen long time a greatest impact in the western to central peninsula. A famine prediction application is modeled using machine learning technique. Predicting the famine for a region early is used to mitigate the vulnerability of the society at risk. Machine learning techniques are experimented on past data collected between 2004 and 2005 in Uganda. The performance of machine learning methods named Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbors (k-NN) and Decision tree classifier in prediction of famine were assessed empirically. SVM and k-NN methods give better result than the rest of the methods, moreover the region of convergence produced by Support Vector Machine can be used by strategic planner in cut-off determination of famine prone management.

II. LITERATURE REVIEW

M.Kalimuthu et.al(2020) "Crop Prediction using Machine Learning"

In general, agriculture is the backbone of India and also plays an important role in Indian economy by providing a certain percentage of domestic product to ensure the food security. But now-a-days, food production and prediction is getting depleted due to unnatural climatic changes, which will adversely affect the economy of farmers by getting a poor yield and also help the farmers to remain less familiar in forecasting the future crops. This research work helps the beginner farmer in such a way to guide them for sowing the reasonable crops by deploying machine learning, one of the advanced technologies in crop prediction. Naive Bayes, a supervised learning algorithm puts forth in the way to achieve it. The seed data of the crops are collected here, with the appropriate parameters like temperature, humidity and moisture content, which helps the crops to achieve a successful growth. In addition, as the software, a mobile application for Android is being developed. The users are encouraged to enter parameters like temperature and their location will be taken automatically in this application in order to start the prediction process.

ZeelDoshi et.al (2018) "Intelligent Crop Recommendation System Using Machine Learning Algorithm."

Agriculture is a major contributor to the Indian economy. The mainstream Indian population depends either explicitly or implicitly on agriculture for their livelihood. It is, thus, irrefutable that agriculture plays a vital role in the country. A vast majority of the Indian farmers believe in depending on their intuition to decide which crop to sow in a particular season. They find comfort in simply following the ancestral farming patterns and norms without realizing the fact that crop output is circumstantial, depending heavily on the present-day weather and soil conditions. However, a single farmer cannot be expected to take into account all the innumerable factors that contribute to crop growth before reaching a consensus about which one to grow. A single misguided or imprudent decision by the farmer can have undesirable ramifications on both himself as well as the agricultural economy of the region. A combination of Big Data Analytics and Machine Learning can effectively help alleviate this issue. In this paper, we present an intelligent system, called Agro Consultant, which intends to assist the Indian farmers in making an informed decision about which crop to grow depending on the sowing season, his farm's geographical location, soil characteristics as well as environmental factors such as temperature and rainfall.

SHILPA MANGESH PANDE et al (2021) "Crop Recommender System Using Machine Learning Approach"

Agriculture and its allied sectors are undoubtedly the largest providers of livelihoods in rural India. The agriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP). Blessing to the country is the overwhelming size of the agricultural sector. However, regrettable is the yield per hectare of crops in comparison to international standards. This is one of the possible causes for a higher suicide rate among marginal farmers in India. This paper proposes a viable and user-friendly yield prediction system for the farmers. The proposed system provides connectivity to farmers via a mobile application. GPS helps to identify the user location. The user provides the area & soil type as input. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. To predict the crop yield, selected Machine Learning algorithms such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN) are used. Among them, the Random Forest showed the best results with 95% accuracy. Additionally, the system also suggests the best time to use the fertilizers to boost up the yield.

Anna Chlingaryanet.al(2018) "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review".

Accurate yield estimation and optimized nitrogen management is essential in agriculture. Remote sensing (RS) systems are being more widely used in building decision support tools for contemporary farming systems to improve yield production and nitrogen management while reducing operating costs and environmental impact. However, RS based approaches require processing of enormous amounts of remotely sensed data from different platforms and, therefore, greater attention is currently being devoted to machine learning (ML) methods. This is due to the capability of machine learning based systems to process a large number of inputs and handle nonlinear tasks. This paper discusses research developments conducted within the last 15 years on machine learning based techniques for accurate crop yield prediction and nitrogen status estimation. The paper concludes that the rapid advances in sensing technologies and ML techniques will provide cost-effective and comprehensive solutions for better crop and environment state estimation and decision making. More targeted application of the sensor platforms and ML techniques, the fusion of different sensor modalities and expert knowledge, and the development of hybrid systems combining different ML and signal processing techniques are all likely to be part of precision agriculture (PA) in the near future.

Rushika Ghadge et.al(2018) "Prediction of Crop Yield using Machine Learning".

Looking at the current situation faced by farmers in Maharashtra, we have observed that there is an increase in suicide rate over the years. The reasons behind this includes weather conditions, debt, family issues and frequent change in Indian government norms. Sometimes farmers are not aware about the crop which suits their soil quality, soil nutrients and soil composition. The work proposes to help farmers check the soil quality depending on the analysis done based on data mining approach. Thus the system focuses on checking the soil quality to predict the crop suitable for cultivation according to their soil type and maximize the crop yield with recommending appropriate fertilizer.

III. CONCLUSION

This method will forecast crop and crop production based on an examination of crop specifics, i.e., agricultural data. This research concludes that farmers, in particular, should develop and apply a system based on machine learning approaches to get more accurate outcomes when making choices. Decision-making is crucial in agriculture for profit or loss, thus it is more necessary to make consistent decisions. Using a random forest approach allows for reliable and efficient crop prediction with no mistakes in the model. Climate characteristics such as average temperature, average humidity, average rainfall, and route map for a chosen crop for certain features with date specification are analysed as concerns that impact crop production. This approach for crop prediction and crop yield prediction uses many machine learning methods such as K-Near Neighbor, Decision Tree, and Random Forest Classifier to boost yield and consequent profit of agricultural output. Certain of the features include algorithm verification and validation, crop forecast utilising some factors, crop analysis, and yield guidance.

REFERENCES

[1] "data.gov.in." [Online]. Available: https://data.gov.in/
[2] Ananthara, M. G., Arunkumar, T., &Hemavathy, R. (2013, February). CRY—an improved crop yield prediction model using bee hive clustering approach for agricultural data sets. In 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (pp. 473-478). IEEE.
[3] Awan, A. M., & Sap, M. N. M. (2006, April). An intelligent system based on kernel methods for crop yield prediction. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 841-846). Springer, Berlin, Heidelberg.
[4] Bang, S., Bishnoi, R., Chauhan, A. S., Dixit, A. K., & Chawla, I. (2019, August). Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1-6). IEEE.
[5] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., &Chaudhari, A. N. (2018, August). Crop Yield Prediction Using Data Analytics and Hybrid Approach. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-5). IEEE.
[6] Gandge, Y. (2017, December). A study on various data mining techniques for crop yield prediction. In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT) (pp. 420-423). IEEE.
[7] Gandhi, N., Petkar, O., & Armstrong, L. J. (2016, July). Rice crop yield prediction using artificial neural.