

Original Article

Supervised Machine Learning for Maize Yield Prediction: A Case Study

Amit Barai ^{1*}, Md Arafat Al Ajmir Sarker ², Ayon Sen ³

1. Computer Science & Engineering, Daffodil International University, Bangladesh.
2. Computer Science & Engineering, Daffodil International University, Bangladesh.
3. Computer Science & Engineering, East West University, Bangladesh.

* Correspondence: amitbarai2001@gmail.com

Abstract: Crop yield forecasting is crucial for agricultural planning, resource allocation, and food security. The use of supervised machine learning algorithms for predicting maize yield with a focus on one of the most significant staple crops globally is examined in this case studies. This study created a precise prediction model using historical weather information, soil characteristics, and management techniques. It begins with data collection from various sources, including weather, soil, windspeed, sunshine, rainfall, temperature, humidity and agricultural records. Data is taken all over Bangladesh. Preprocessing techniques are then applied to handle missing values, normalize the data, and engineer relevant features. Feature selection methods are employed to identify the most influential factors for maize yield prediction. Next, a range of supervised machine learning algorithms is considered, and the most suitable models are chosen based on their performance and relevance to the problem at hand. The selected models such as Linear Regression, XGBR, GBR, RFR, DTR are trained using the preprocessed data and evaluated using appropriate metrics to assess their predictive capabilities. The case study identified the key factors that significantly impact maize yields. By applying feature selection techniques, the study seeks to determine the most relevant variables and their respective importance in yield prediction. This outcome provides valuable insights into the factors that should be prioritized for effective yield management and crop optimization. The results and discussions provide insights into the performance of different models and the interpretation of predictive features. Where the highest accuracy is 93.85% for GBR. The challenges faced during the process, such as data limitations and potential biases, and explores opportunities for further research and improvement are also highlighted. Overall, this case study demonstrates the potential of supervised machine learning in accurately predicting maize yields, thereby aiding agricultural planning and decision-making for improved productivity and food security.

Keywords: Maize, Yield, Prediction, GBR

1. INTRODUCTION

After rice and wheat, maize is the most significant cereal crop in Bangladesh. It might provide nutritional benefits. According to a study [1] 100 grams of mature maize seeds contain 9.42 grams of protein, 74.26 grams of carbs, 0.64 grams of sugar, 7.3 grams of dietary fiber, and 365 calories. In addition to improving renal and bone health, maize also controls heart rate and inhibit stomach acidity and causes constipation. Maize also lowers LDL cholesterol and protects against cardiovascular problems, diabetes, and hypertension [2]. Maize is a vital staple crop worldwide, serving as a significant source of food, feed, and industrial raw materials. At present, the annual demand for maize in Bangladesh is around two million tons, but production is 4,700 thousand tones which is a big gap between demand and production [3]. So, accurate yield prediction can assist farmers, policymakers, and stakeholders in making informed decisions regarding crop planning, resource allocation, and market forecasting [4].

This case study aims to harness the power of supervised machine learning techniques to enhance maize yield prediction accuracy, thereby improving agricultural productivity and food security [5]. In recent years, there have been remarkable advancements in computing power, data availability, and machine learning algorithms. These advancements have made it possible to analyze vast amounts of data and develop sophisticated models for crop yield prediction [6]. By leveraging these technologies, this case study aims to explore the potential of supervised machine learning in accurately predicting maize yields, thereby enabling farmers to optimize their cultivation practices and enhance overall productivity [7]. The availability of comprehensive datasets encompassing weather conditions, soil attributes, and management practices presents an opportunity to leverage data-driven approaches for yield prediction [8]. By employing supervised machine learning, this case study seeks to extract meaningful patterns and relationships from the data, enabling stakeholders to make informed decisions based on quantitative predictions rather than relying solely on traditional methods or subjective judgments [9,10]. Huge sums of money are spent on importing maize seeds and products in order to meet the demand [11]. The use of maize in Bangladesh for human, animal, and agricultural Future increases in poultry feed across all categories are planned [12,13]. Additionally, maize has a promising future in Bangladesh because the country's average annual weather has benefited maize cultivation. In Bangladesh, maize can grow well in every environment due to its extensive genetic diversity. In Bangladesh, it often grows year-round and exhibits potential productivity. Few reports revealed that maize is more profitable crop than rice [14]. Reported that maize is more economically profitable than that [15,16] reported maize production has a much higher profit than wheat. Although the rice-based green revolution technology in Bangladesh is a common practice, there is an urgent need to crop diversification due to sustaining its growth. The government of Bangladesh is also seeking to diversify crops to other cereals than rice. The rate of adoption and sustainability of any crop depends upon its economic profitability [17]. Economic profitability is one of the important criteria for assessing the suitability of a new crop technology considering this point maize will be the best option for crop diversification [18]. Regarding the overall backdrop, the present study provides insights into the potential of using machine learning techniques to improve the accuracy of crop yield predictions in Bangladesh, which can aid in crop management and policy-making decisions [19]. Improving the accuracy of maize yield prediction can have significant impacts on agricultural practices, resource allocation, and food security. Enhanced prediction models can aid in mitigating the risks associated with climate change, optimizing resource usage, and supporting sustainable farming practices [20,21]. By demonstrating the effectiveness of supervised machine learning in this domain, this case study aims, to contribute to the broader goal of advancing precision agriculture and ensuring a stable and resilient food supply. The agricultural importance of maize, the advancements in technology and machine learning, the potential for data-driven decision making, and the desire to have a positive impact on agricultural productivity and food security. Custom real-time field dataset; Gradient Boosting Regressor, Linear Regressor, XG Boost Regressor, Random Forest Regressor, and Decision Tree Regressor are some of the methods that are used. This work expands the field of research.

2. MATERIALS AND METHODS

The study subjected to the case study on supervised machine learning for maize yield prediction is the prediction of maize yields based on historical weather data, soil attributes, and management practices. The study aims to develop accurate predictive models that can estimate maize yields with a high level of accuracy. To conduct the research, various instruments and tools will be utilized. Here are some key elements of instrumentation for the case study, Instruments for collecting relevant data sources include weather stations for gathering historical weather data, soil sensors or surveys for obtaining soil attributes, and agricultural records for capturing management practices such as fertilization, irrigation, and crop rotation. Preprocessing techniques is applied to handle missing values, normalize data, and engineer features. Tools like Python libraries (e.g., pandas, NumPy) and data manipulation tools (e.g., Excel) is used for data preprocessing. Various machine learning libraries and algorithms is employed for model development and evaluation. Popular libraries such as scikit-learn, TensorFlow, matplotlib

provide a wide range of algorithms linear regression, decision trees, random forests, XGBR, GBR is used for supervised learning tasks. Programming languages such as Python is used for implementing the machine learning models, data preprocessing, and analysis. These languages offer rich libraries and frameworks for machine learning and data analysis tasks. For evaluation metrics such as accuracy, mean squared error (MSE), root mean squared error (RMSE), and R-squared is used to assess the performance of the developed models. Statistical analysis tools like Python or R can be utilized for calculating these metrics. For visualizing the data, model performance, and interpretability of results is important. Tools like Matplotlib, Seaborn, or Tableau can be used to create visualizations and graphs that aid in understanding the relationships between variables and presenting the findings. For Collaboration tool Google Drive is used for version control, sharing code, and collaboration among researchers. For documentation tool Microsoft Word can be used for writing the research report and paper. The above-mentioned tools and instruments provide a general framework for conducting the case study on supervised machine learning for maize yield prediction. The data collection procedure for real-time maize production on Bangladeshi climate data using supervised learning can be broken down into the following steps: The first step in data collection is to identify the relevant variables that are required to predict maize production. This may include climate variables such as temperature, rainfall, humidity, wind speed, and solar radiation, as well as maize production variables such as the area under cultivation, yield per unit area, and total production. Once the relevant variables have been identified, the next step is to identify the data sources. Climate data can be obtained from various sources such as the Bangladesh Meteorological Department, and maize production data can be obtained from the Bangladesh Bureau of Statistics or from individual maize farmers. The data can be collected manually by visiting the relevant sources or through automated data collection methods such as remote sensing or IoT sensors. The data collection should be done in real-time to ensure that the predictive models are up-to-date. The collected data should be cleaned to remove any missing or erroneous values and preprocessed to ensure that all variables are on the same scale. The data should also be split into training and testing sets. The collected and preprocessed data should be stored in a database or other storage facility for easy access and management. Real-time data collection and monitoring should be maintained to ensure that the predictive models remain up-to-date and accurate. This may involve the use of automated data collection and processing methods. In conclusion, the data collection procedure for real-time maize production on Bangladeshi climate data using supervised learning involves identifying relevant variables, identifying data sources, data collection, data cleaning and preprocessing, storage and management, and continuous monitoring. A few steps must be followed in order to develop an algorithmic prediction of maize production. The entire system for predicting maize yield is fully explained in this section. Collecting the required data is the most challenging aspect of a research investigation. Before using the model, prepare the dataset. A custom dataset needs to be created first. Creating a dataset and employing models are the challenging aspects of the research process. Selecting the model that best fits the dataset could be difficult. This section went over the methodology's full approach.

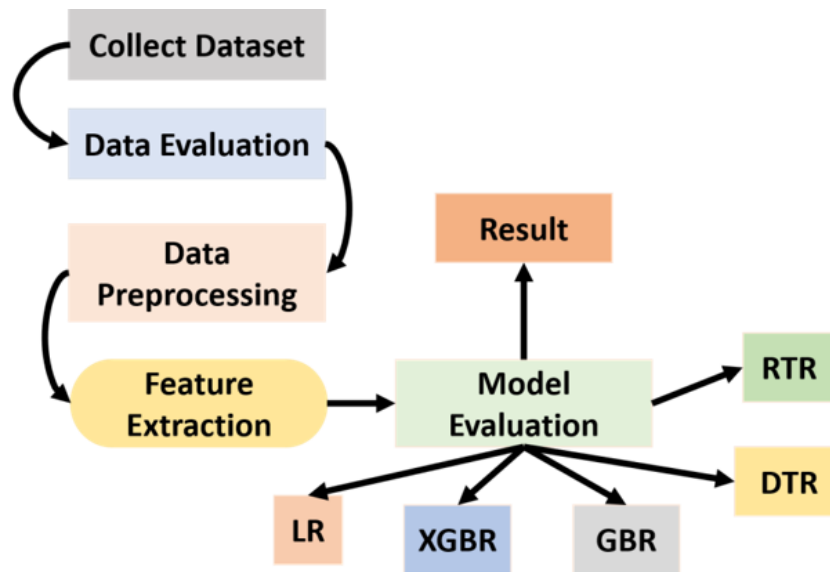


Figure 01: Process of ML Model for maize prediction

The problem must be identified before moving on to the first and most crucial phase. Both the input and output variables have been selected. The desired outcome of the detection is displayed in the output variable. In our model, we contrasted the output of various ML methods. The data collection, planning, and implementation were the parts of our study paper that proved to be the most challenging. Having goals and resources was also essential for creating our strategy. The data required to produce this dataset was given by the Bangladesh Bureau of Statistics (BBS), collected data between 1968 and 2021 on maize yields from 23 locations of Bangladesh over a 54-year period. The eleven parameters include location, bale productivity, hectare area, year, windspeed, sunshine, rainfall, lowest and maximum temperatures, humidity, and cloud cover.

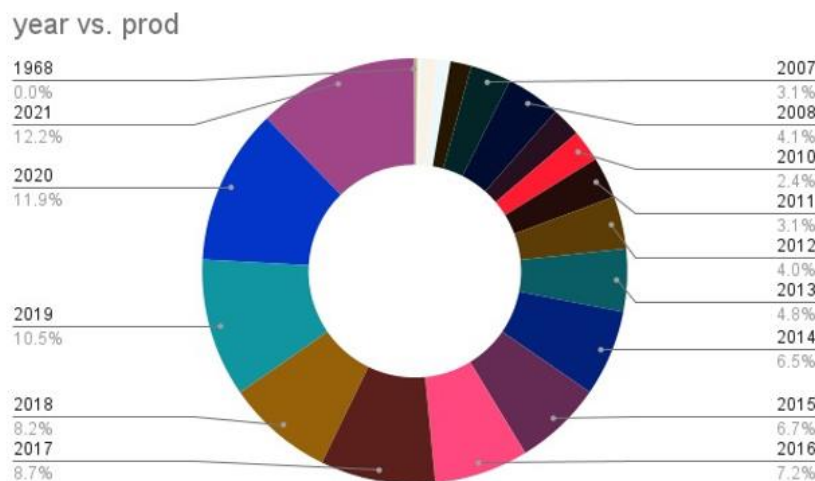


Figure 02: Comparison of production variation by year and production

This graph illustrates how the value of production changes with the year and prod. The outcomes of a statistical study were analyzed using the Google Collaboratory platform. Computers now have additional space as a result of the extra GPU and TPU that Google Collaboratory added. Python was the programming language used. Convert the PDF data into Excel before you begin gathering data. The parameter labels for some of the variables were altered after the gathering of our data. When the dataset was initially analyzed, numerous significant inconsistencies were found using the variables

"district," "area in hectares," "production in bales," and "minimum temperature." The abbreviated form was used to manually alter a variety of variables, including "district" to "region," "Area in Hectare" to "area," "Production in Bales" to "production bale," and "Minimum Temperature" to "mintemp" and "Maximum Temperature" to "max temp." It was done to preprocess the variables. The missing value is then discovered after preprocessing the collected data. Any missing values in our data sets were not discovered. After that, the custom dataset was machine-fit. Making a decision about the test and training method is the most important step in any research project. Partition our dataset into training and testing. The model was trained on 813 of the 1161 data, and it was tested on 348 of the 1161 data. Therefore, allocate 30% of the resources to testing and 70% to training. Linear regression is a machine learning algorithm for supervised learning. Analyses of regression are carried out. Using independent variables, regression creates a value for the objective prediction. Prediction is the main use of finding relationships between variables. The XGB Regressor usually ranks each feature used for prediction in order of relevance. Gradient boosting has the advantage that, after boosted trees have been built, obtaining relevance ratings for each attribute is not too difficult. Employing gradient boosting for regression. This estimator permits the forward stage-wise optimization of any differentiable loss function while building an additive model. When the specified loss function has a negative gradient at each level, a regression tree is fitted. In classification and regression issues, Random Forest and other supervised machine learning algorithms are commonly utilized. It creates decision trees from several samples, sorts them based on their average, and regresses them in accordance with a consensus judgment. Train Decision trees, a type of supervised machine learning technique, are used by the Auto ML tool to categorize or regress the data using yes or no answers to particular questions. The quality and precision of each categorization system determines the outcomes. The accuracy of machine learning models is used in this study to validate them. How much of the variance of an independent variable can be explained by a dependent variable is determined by the R squared statistic. All of the ML methods employed are contrasted in this section. The algorithm error rate graph's data MSE, MAE, and RMSE.

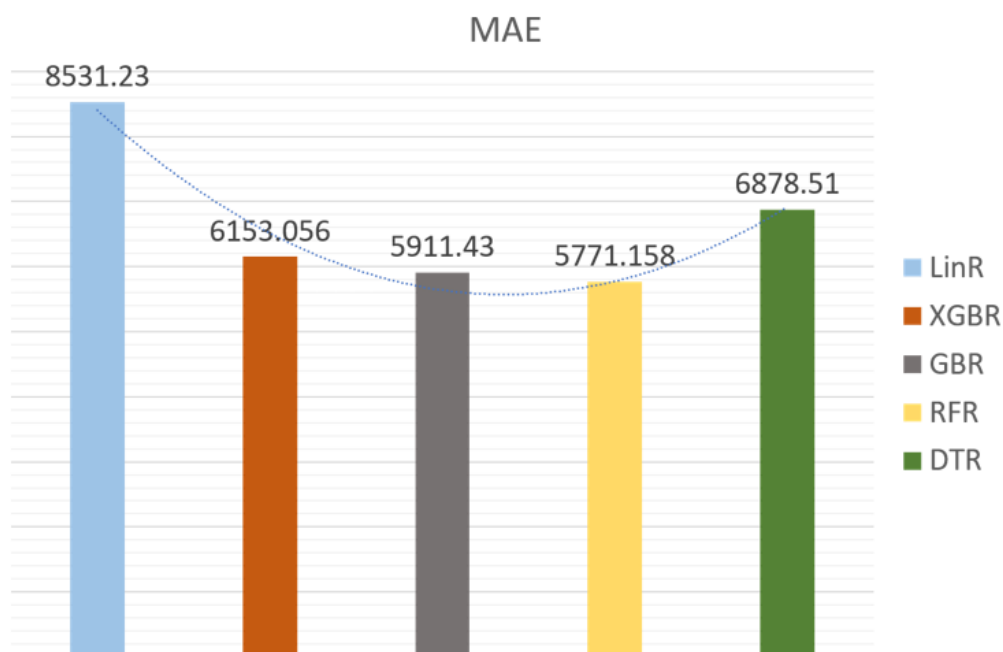


Figure 03: Displaying algorithmic characteristics for MAE

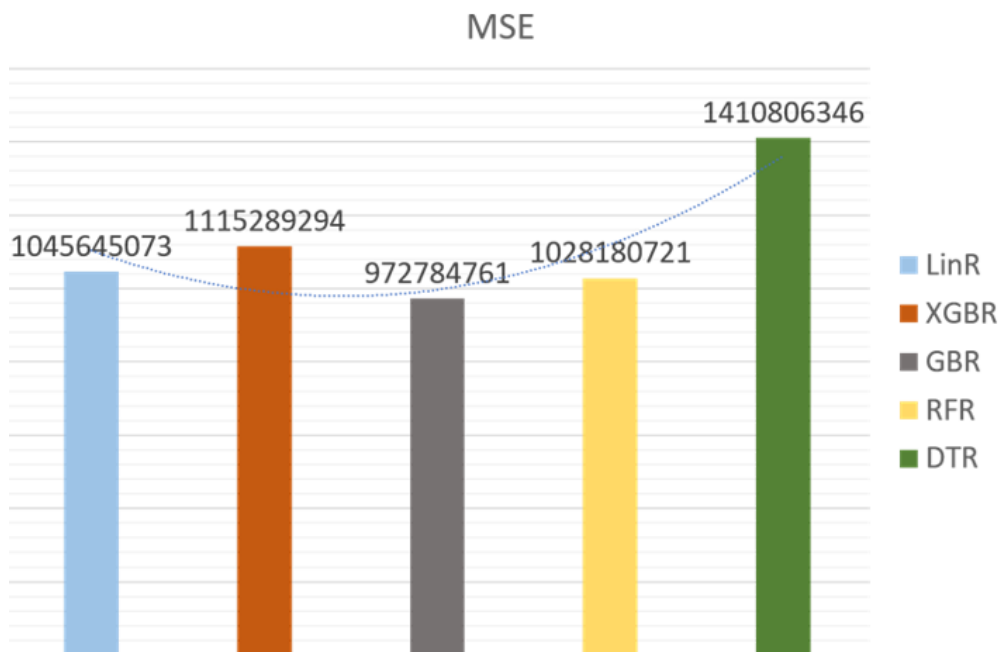


Figure 04: Displaying algorithmic characteristics for MSE

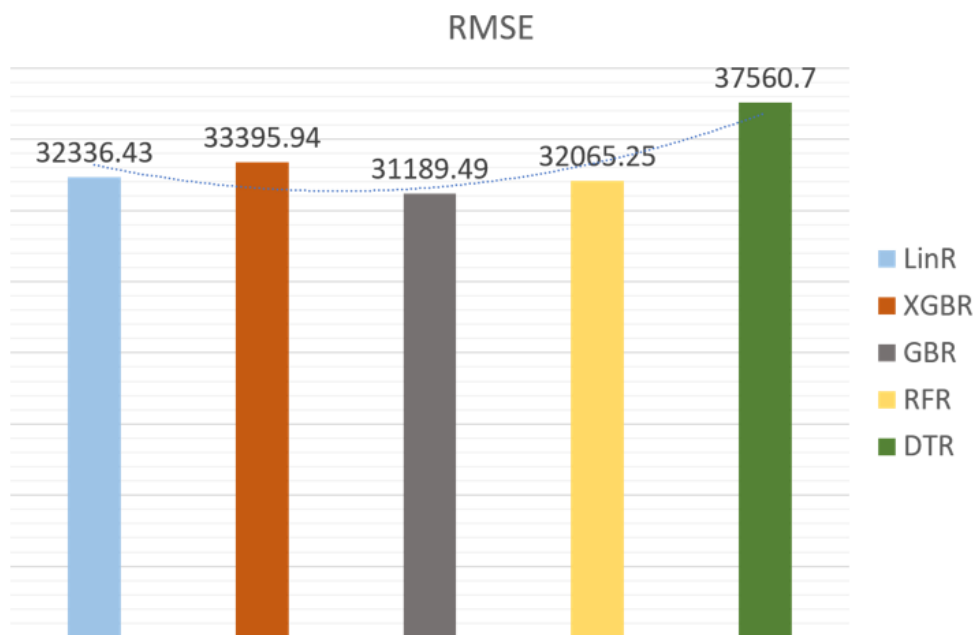


Figure 05: Displaying algorithmic characteristics for RMSE

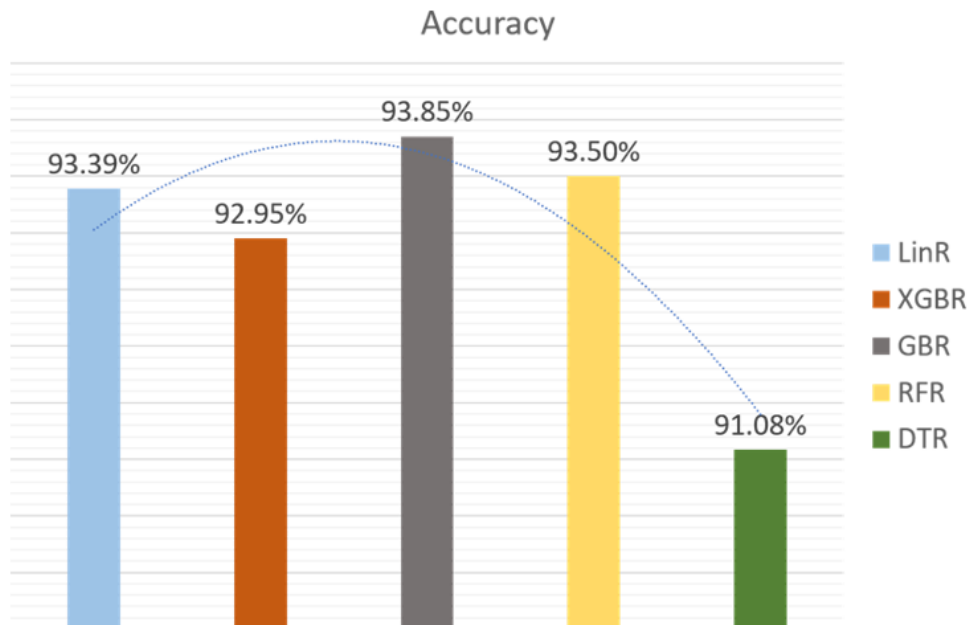


Figure 06: Displaying algorithmic characteristics for Accuracy

The following figure 03,04,05,06 showing the algorithmic characteristics for different error rate of maize prediction. Whereas the highest MAE 8531.23 Linear Regressor and the lowest value of MAE 5771.158 for Random Forest Regressor but Decision Tree Regressor is the highest value of MSE is 1410806346.09 and lowest value of Gradient Boosting Regressor is 972784760.98. Same as MES, the lowest value of RMSE the Gradient Boosting Regressor is 31189.49 and the highest value of Decision Tree Regressor is 37560.70. On the other hand, the highest accuracy is 93.85% for Gradient Boosting Regressor and 91.08% accuracy for Decision Tree Regressor as the lowest.

3. RESULT & DISCUSSION

Use a dataset with 1161 total data points, each of which includes the region, area, product, year, cloud cover, humidity, maximum temperature, minimum temperature, rainfall, sunshine, and wind speed. It used various machine learning models, including Random Forest Regressor, Gradient Boosting Regression, Linear Regression, and XG. Each model assessed accuracy like train set accuracy and test set accuracy as well as MAE, MSE, RMSE, and R2. Results showed that Linear Regression (93.39%), XGBR (92.95%), Gradient Boosting (93.85%), Decision Tree (91.08%), and Random Forest (93.50%) were all accurate predictors.

TABLE 01: Measurement of algorithmic inaccuracy

Algorithm	MAE	MSE	RMSE
XG Boost Regressor	6153.05	1115289294.16	33395.94
Linear Regressor	8531.23	1045645072.93	32336.43
GradientBoosting Regressor	5911.42	972784760.98	31189.49
RandomForest Regressor	5771.15	1028180721.41	32065.25
Decision Tree Regressor	6878.50	1410806346.09	37560.70

This is explained in the table 01 the highest MAE 8531.23 Linear Regressor and the lowest value of MAE 5771.158 for Random Forest Regressor but Decision Tree Regressor is the highest value of MSE is 1410806346.09 and lowest value of Gradient Boosting Regressor is 972784760.98. Same as MES, the lowest value of RMSE the Gradient Boosting Regressor is 31189.49 and the highest value of Decision Tree Regressor is 37560.70. On the other hand, the highest accuracy is 93.85% for Gradient Boosting Regressor and 91.08% accuracy for Decision Tree Regressor as the lowest. Accuracy, flexibility, and extensibility were the three guiding principles used to develop the machine learning baseline. Focus first on the accuracy of correctly applying machine learning and delivering comprehensible characteristics. When working with time series data, such as maize yield, features developed using values from previous years are used, such as the maize yield trend. When data from previous years are merged with features, more effort has been made to avoid information leaking. The findings of a supervised learning analysis of maize production on Bangladeshi climate data will depend on the specific variables and models used in the analysis. However, some findings that could be of interest include:

- The most important climate factors for maize, including temperature, rainfall, humidity, and soil moisture, have been identified.
- The prediction of maize yield based on climate variables for different regions and seasons in Bangladesh.
- Evaluating the precision with which various supervised learning algorithms can forecast maize yield using climate data.
- Identification of the optimal range of climate variables for maize production in Bangladesh and the potential impact of climate change on maize yields.
- Assessment of how well various agricultural management techniques such as irrigation and insect control affect maize yields under various climate conditions.
- The discovery of additional variables that might influence maize output, such as crop rotation, soil type, and use of fertilizer, which could be taken into account in future investigations.

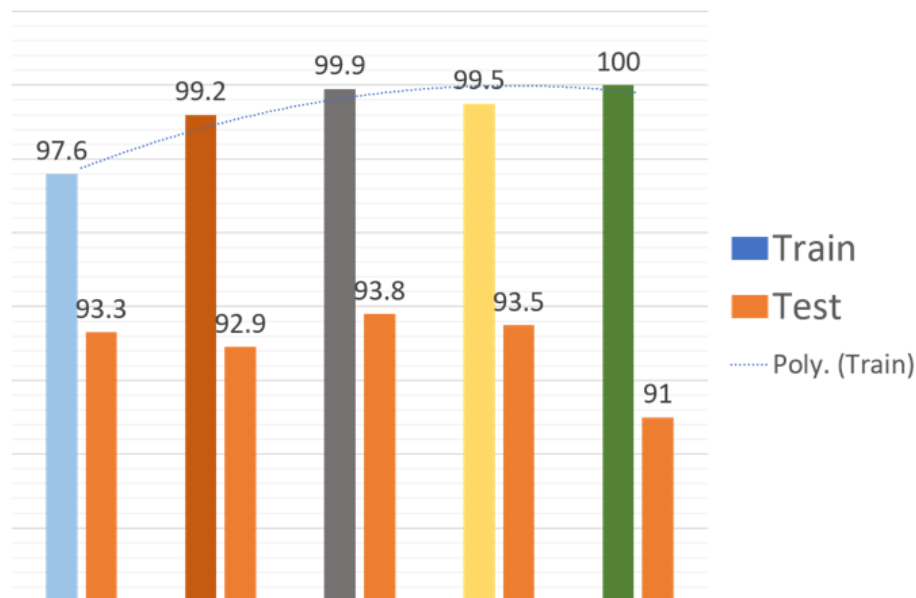


Figure 07: Algorithmic comparison on Train and Test Accuracy

In general, the results of a supervised learning analysis of maize output on climate data from Bangladesh can offer useful insights into the variables that influence crop yields and guide decision-making for farmers, policymakers, and other stakeholders in the region's agriculture. Schulthess et. al developed a methodology using the Hybrid Maize model and satellite imagery to create a yield gap map for maize in northwestern Bangladesh. Where the map helps identify constraints in farmers' fields with significant yield gaps and enables the generation of optimized crop management recommendations based on

high-yielding farmers' data. The average potential yield was estimated at 12.87 Mg/ha in this high-yielding environment. Choudhury et. al conducted in Bangladesh determined the optimum sowing window for maize in the northern and western regions. The model accurately forecasted yields 45 days prior to harvest, providing valuable information for farmers and policymakers to meet future maize demands. Hossain et. al introduced WPSRY (Weather-based Prediction System for Rice Yield), an approach for forecasting rice yield in Bangladesh. WPSRY combines Neural Networks to predict weather parameters and Support Vector Regression to estimate rice yields, resulting in accurate predictions and promising accuracy. This system is crucial for low-lying countries like Bangladesh that are vulnerable to climate change's impact on agriculture. Kalaiarasi et. al proposed Multi-parametric Deep Neural Network (MDNN) utilizes Growing-Degree Day (GDD) as a measure to model the impact of climate changes on crop yield prediction. and considering multiple parameters related to weather and soil, MDNN achieves a mean accuracy of 91.84% for predicting crop yield, outperforming traditional Deep Neural Network (DNN) approaches which demonstrated the enhanced accuracy of MDNN by incorporating climate, weather, and soil parameters in the analysis. Nishant et al. 33 states, 2.5 lakh observations, and more than 20 crops were used in this article. The Indian Government Repository is the data's source. Here, the algorithms Lasso, ENet, Kernel Ridge, Root Mean Square Error, and Stacked Regression are utilized. Better precision can be found in Kernel Ridge. Islam, et al. used rice, Boro rice, Aman rice, potato, jute, wheat, and so forth. Data for the websites BBS and BMD were gathered from SRDI, 70 books, and BARC. Information from zone number 28, covers Narshingdi, Dhaka, Narayanganj, Gazipur, Mymensingh, and Tangail. It is possible to forecast agricultural yield using DNN, Logistic Regression, SVM, and RF, and DNN consistently has the greatest accuracy. Kumar et al. They have gathered 3101 observations from the Kaggle repository. The best precision may be achieved using RF when compared to DT and SVR. Champaneri, Mayank, et al. used 10 years' worth of annual crop abstracts to research every district in Maharashtra. The accuracy of RF has been employed, and it is 75%. They also created a webpage. Gandhi, Niketa, et al. Around four years' worth of records and 27 Maharashtra districts' worth of data were gathered. It has been used to apply MAE, RMSE, RAE, SMO, WEKA, and SVM. More effectively than SMO were Nave Bayes, Bayes Net, and Multilayer Perceptron. Kale, Shivani S. This study focuses on Maharashtra. The Indian Government Website gathered data from 1997 to 2014. There has been an application of ANN with forward and backward propagation, MSE, linear regression, MAE, and RMSE. It is 82% accurate. Pandith, Vaishali, et al. The Jammu Department of Agriculture's Talab Tillo was consulted for information while they conducted a study on mustard. With 11 input parameters, the data has 5000 cases. When KNN and RNN are employed, the results are better than when MLR, NB, RF, and ANN are utilized. Bondre, Devdatta A. et al. This is an Indian research article that compiled data from various books and websites over five years. Here, RF and SVM are employed. SVM yields about 99.47%, whereas RF yields 97.48%. In Madhya Pradesh, India, Veenadhari, S. et al. studied wheat, maize, paddy, and soybeans. They utilized 20 years' worth of climatic information from five districts in Madhya Pradesh, including rainfall, the greatest and lowest temperatures, the potential for evapotranspiration, cloud cover, and the frequency of rainy days. Researchers utilized C4.5 with a 75% accuracy rate, and the C# programming language was used to build a web application on the .net framework. In Bangladesh, Ahamed, AT M. Shakil, et al. performed research on wheat, potatoes, AMON, AUS, and BORO. From 2009 to 2011, they gathered data from BARI (Bangladesh Agricultural Research Institute) from 15 districts in Bangladesh, including low and high temperatures, area irrigated, humidity and rainfall across all districts, and area cultivated across all crops taken into consideration throughout the districts. Accuracy is not specified and the usage of linear regression, k-NN, neural networks, and ANN has been made. In Munshi Ganj, Bangladesh, Mahdi, Mostafa Didar, et al. do potato research. They gathered information from the websites of Google Earth, CHIRPS (Climate Hazards Group Infrared Precipitation with Station data), and USGS Earth Explorer. The RMSE and validation error for the rainfall analysis of the time series using an LSTM with the Deep Gaussian process were both 0.433. Mamun et al. focused on estimating jute yield in Bangladesh using machine learning techniques. The research compares various algorithms and finds that the decision tree regressor performs the best, achieving an average prediction accuracy of 96%. The findings of this

study are expected to be valuable for future researchers working on jute cultivation. Haque et al. Used the ARIMA model methodology, the experiment reveals a positive trend in the crop area, production, and yield of wheat in Bangladesh, which has significant implications for the country's economy, employment, and food security. Nigam et al. utilizing various machine learning techniques and comparing their performance based on mean absolute error, the study aims to assist farmers in making informed decisions about crop selection, considering factors such as temperature, rainfall, and area. This has the potential to improve income prospects for farmers and transform the agricultural sector.

4. CONCLUSION

Aiming to use climatic data and supervised learning methods to create a real-time approach for predicting maize production in Bangladesh. The main points and conclusions of our study are highlighted in this discussion part, along with the effectiveness of various supervised learning algorithms and pertinent issues. Firstly, various supervised learning algorithms, including Linear Regressor, XG Boost Regressor, Gradient Boosting Regressor and Random Forest Regressor, to predict maize production based on climate variables were examined. Each model assessed accuracy like train set accuracy and test set accuracy as well as MAE, MSE, RMSE, and R2. Results showed that Linear Regression (93.39%), XGBR (92.95%), Gradient Boosting (93.85%), Decision Tree (91.08%), and Random Forest (93.50%) were all accurate predictors. These algorithms' performance evaluations indicated variances in their forecast accuracy and capacity to grasp the intricate connections between climatic variables and maize yield. This investigation gave insightful information about the applicability of several algorithms to this particular prediction problem. We also addressed the issue of data shortage in several areas of Bangladesh in our study. The accuracy and generalizability of the forecasting model may be hampered in some regions by the lack of widespread access to extensive climate data. The researchers underscored the necessity for additional data gathering and the significance of setting up dependable climate monitoring networks across the nation to lessen this limitation. The forecasting model would be more reliable if climatic data were more readily available and of higher quality. In conclusion, a real-time method for predicting maize production in Bangladesh based on climate data using supervised learning techniques is presented. The findings demonstrated the potential of supervised learning algorithms in accurately predicting maize production by leveraging climate variables. The research outcomes highlighted the performance variations among different supervised learning algorithms and their suitability for predicting maize production. This knowledge can guide future studies and assist stakeholders in selecting the most appropriate algorithm for their specific needs. Moreover, we emphasized the importance of selecting relevant input features and addressing data scarcity issues. By considering the influential climate variables and promoting data collection efforts, the accuracy and applicability of the predictive model can be enhanced.

5. RECOMMENDATIONS

In Bangladesh, farming is a viable sector that may be used to cultivate maize effectively by preserving soil health, enhancing water conservation, applying cultivation technology, and raising farmers' income. This case study is vast and offers opportunities for advancements in various aspects. Here are some potential areas for further exploration, Expand the scope of data used for maize yield prediction by integrating multiple data sources. This can include incorporating additional environmental data (e.g., satellite imagery, climate data), soil characteristics, management practices, and socioeconomic factors. Explore techniques to handle heterogeneous data and effectively fuse different data types to improve prediction accuracy. Investigate methods to scale the supervised machine learning models for maize yield prediction across larger regions or different cropping systems. Develop techniques to transfer knowledge and models from one region to another, accounting for variations in soil, climate, and management practices. Consider the challenges of data heterogeneity, model adaptability, and localization. Extend the scope of the case study to include crop management recommendations based on the predicted maize yields. Develop decision support systems that provide actionable insights to farmers, including optimal planting dates, irrigation scheduling, fertilizer application rates, and pest

management strategies. Consider the integration of precision agriculture technologies to facilitate site-specific recommendations. Extend the case study to focus on long-term yield forecasting, beyond a single growing season. Investigate the potential of supervised machine learning to predict yield trends over multiple years, accounting for interannual variability and climate change impacts. Consider the integration of climate projection models and predictive analytics to provide insights into future yield patterns. These areas provide a starting point for further developments in supervised machine learning for maize yield prediction. Continued research, collaboration, and innovation in these directions can contribute to the advancement of agricultural analytics, support sustainable farming practices, and enhance global food security.

REFERENCES

- [1] Schulthess, U., et al. "Mapping field-scale yield gaps for maize: An example from Bangladesh." *Field Crops Research* 143 (2013): 151-156.
- [2] Choudhury, Apurba Kanti, et al. "Optimum Sowing Window and Yield Forecasting for Maize in Northern and Western Bangladesh Using CERES Maize Model." *Agronomy* 11.4 (2021): 635.
- [3] Hossain, Md Akter, et al. "Predicting rice yield for Bangladesh by exploiting weather conditions." 2017 international conference on information and communication technology convergence (ICTC). IEEE, 2017.
- [4] Kalaiarasi, E., and A. Anbarasi. "Crop yield prediction using multi-parametric deep neural networks." *Indian Journal of Science and Technology* 14.2 (2021): 131-140.
- [5] Nishant, Potnuru Sai, et al. "Crop yield prediction based on Indian agriculture using machine learning." 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.
- [6] Islam, Tanhim, Tanjir Alam Chisty, and Amitabha Chakrabarty. "A deep neural network approach for crop selection and yield prediction in Bangladesh." 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). IEEE, 2018.
- [7] Kumar, Y. Jeevan Nagendra, et al. "Supervised machine learning approach for crop yield prediction in agriculture sector." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.
- [8] Champaneri, Mayank, et al. "Crop yield prediction using machine learning." *Technology* 9 (2016): 38.
- [9] Gandhi, Niketa, et al. "Rice crop yield prediction in India using support vector machines." 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE, 2016.
- [10] Kale, Shivani S., and Preeti S. Patil. "A machine learning approach to predict crop yield and success rate." 2019 IEEE Pune Section International Conference (PuneCon). IEEE, 2019.
- [11] Pandith, Vaishali, et al. "Performance evaluation of machine learning techniques for mustard crop yield prediction from soil analysis." *Journal of scientific research* 64.2 (2020): 394-398.
- [12]
- [13] Bondre, Devdatta A., and Santosh Mahagaonkar. "Prediction of crop yield and fertilizer recommendation using machine learning algorithms." *International Journal of Engineering Applied Sciences and Technology* 4.5 (2019): 371-376.
- [14] Veenadhari, S., Bharat Misra, and C. D. Singh. "Machine learning approach for forecasting crop yield based on climatic parameters." 2014 International Conference on Computer Communication and Informatics. IEEE, 2014.
- [15] Ahamed, AT M. Shakil, et al. "Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh." 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). IEEE, 2015.
- [16] Mahdi, Mostafa Didar, et al. "A Deep Gaussian Process for Forecasting Crop Yield and Time Series Analysis of Precipitation Based in Munshiganj, Bangladesh." IGARSS 2020- 2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2020.

-
- [17] Mamun, Shahriar, et al. "JuteBangla: A Comparative Study on Jute Yield Prediction using Supervised Machine Learning Approach based on Bangladesh Perspective." 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2022.
- [18] Haque, Kohinoor, Md Khairul Islam, and Abdus Sattar. "Wheat Production Forecasting in Bangladesh Using Deep Learning Techniques." 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2022.
- [19] Nigam, Aruvansh, et al. "Crop yield prediction using machine learning algorithms." 2019 Fifth International Conference on Image Information Processing (ICIIP). IEEE, 2019.
- [20] Medar, Ramesh, Vijay S. Rajpurohit, and Shweta Shweta. "Crop yield prediction using machine learning techniques." 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). IEEE, 2019.
- [21] Reddy, D. Jayanarayana, and M. Rudra Kumar. "Crop yield prediction using machine learning algorithm." 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2021.
- [21] Researchgate, https://www.researchgate.net/publication/363282633_A_Brief_Review_on_the_Present_Status_Problems_And_Prospects_of_Maize_Production_in_Bangladesh, 12:31,5/28/2023.