IJSES
International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 10-16, September 2026.
ISSN (Online): 2456-7361

Machine Learning-Based Cotton Production Prediction Using Climatic and Agricultural Data

Jashanpreet Singh[1], Dr. Lal Chand[2]
[1]Student, Department of Computer Science, Punjabi University, Patiala, Punjab, India
[2]Assistant Professor, Department of Computer Science, Punjabi University, Patiala, Punjab, India
Email address: [1]jashanpreetsinghakku@gmail.com, [2]lc.panwar5876@gmail.com

Abstract—*Cotton production is one of the important aspects of agriculture in various nations. Crop yield prediction is crucial for farmers since it enables them to make better judgments regarding crop farming. In this study, a machine learning-based methodology is proposed for predicting cotton production based on historical agricultural and climatic data. The proposed methodology is based on historical agricultural and climatic data, and various important parameters such as area, rainfall, temperature, humidity, fertilizer, and district information are included in the dataset, collected from various regions and years. The data preprocessing technique is used to prepare the data for training the model. In this study, three machine learning algorithms, namely Random Forest, Gradient Boosting, and Extreme Gradient Boosting (XGBoost), were used to predict cotton production. The proposed model is evaluated using various regression metrics such as (R^2) score, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The experimental results show that the XGBoost model performed better than other machine learning algorithms, achieving a score of 0.9664, followed by Gradient Boosting and Random Forest. The results show that machine learning algorithms can effectively predict crop yield based on climatic parameters and can be used to enhance crop production forecasting.*

Keyword—*Cotton Yield Prediction, Machine Learning, Random Forest, XGBoost, Precision Agriculture, Climate Data, Agricultural Data, Crop Yield Forecasting.*

I. Introduction

Cotton is one of the most important cash crops in the world and plays a vital role in the economic and agricultural scenario of developing countries [1]. In countries like Pakistan and other regions of the world where cotton is grown, the cash crop plays a significant role in boosting agricultural productivity and providing employment to the local population [2]. Cotton not only contributes to the economic activities of the nation but also contributes significantly to the export earnings of the country [5]. Due to its importance in the economic scenario of the nation, improving the yield prediction and management of the cash crop has become a vital area of research in modern agriculture [6].

The climate variability and changes in the environment have become significant factors in affecting the yield and productivity of cash crops like cotton. The increase in temperature and changes in rainfall patterns have become significant contributors to climate variability and changes in the environment [7]. Climate changes have been recognized as one of the most important factors that can adversely affect the sustainability of agriculture in the future [10]. Cotton is one of the climate-sensitive cash crops that can be adversely affected by climate changes. Climate changes can have significant impacts on the yield and productivity of cash crops like cotton. In regions like Punjab, which are highly dependent on agriculture for economic activities, climate-related risks have become significant challenges [8].

Therefore, understanding the link between agricultural production and climatic factors has become a crucial component of agricultural research. Statistical analysis and econometric models have been used to study crop yield and its trends. However, these models often fail to cope with nonlinear relationships and complex interactions of environmental variables with crop growth patterns [9]. As agricultural systems become increasingly complex due to climate variability and other factors, there is an urgent need for sophisticated models that can cope with complex agricultural data.

Machine learning (ML) has proven to be an effective technique for dealing with agricultural challenges. ML techniques can process complex heterogeneous data and generate accurate predictions for crop yield forecasting [11]. Recent studies have shown that ML-based models can be used to improve agricultural prediction problems compared to traditional statistical analysis techniques. ML-based models can cope with nonlinear relationships and complex interactions of environmental variables with crop growth patterns [3]. ML-based models have been effectively utilized in agricultural challenges such as crop yield prediction, irrigation management, climate impact analysis, and disease identification.

Thus, incorporating climate variables and machine learning techniques can improve the accuracy of crop yield forecasting models for cotton production. Climate variables such as temperature, rainfall, humidity, and soil conditions are integral factors that influence crop growth patterns and ultimately determine crop yield. Machine learning models can be used to analyze historical data on climate variables and crop production to understand complex relationships and predict crop yield for future crop production [4]. This is extremely useful for farmers and agricultural planners who need to take effective decisions regarding crop production.

Recent developments in smart agriculture have once again highlighted the need for data-driven strategies to boost crop productivity. Machine learning strategies have provided researchers with the opportunity to create predictive models that can aid precision farming strategies and improve decisionmaking mechanisms in agriculture. This helps farmers to forecast yield fluctuations, optimize fertilizers, and reduce

Volume 10, Issue 9, pp. 10-16, September 2026. ISSN (Online): 2456-7361

the adverse effects of climate change. Therefore, ML-based models for agriculture forecasting are becoming increasingly important for modern-day sustainable agriculture.

Inspired by these advances, this study is motivated to predict cotton production using various machine learning methods on historical agricultural and climatic data. The ultimate aim is to build precise predictive models that can effectively handle intricate relationships between climatic conditions and cotton production. Various machine learning methods will be tested to assess their predictive capability and potential for cotton yield forecasting. By harnessing historical production records and climatic conditions, this study will contribute to informed decision-making and improved crop productivity forecasting in agriculture.

The overall framework of the proposed cotton yield prediction system is illustrated in Fig. 1. The framework highlights the integration of climate variables, agricultural data, and machine learning algorithms to estimate cotton production patterns.

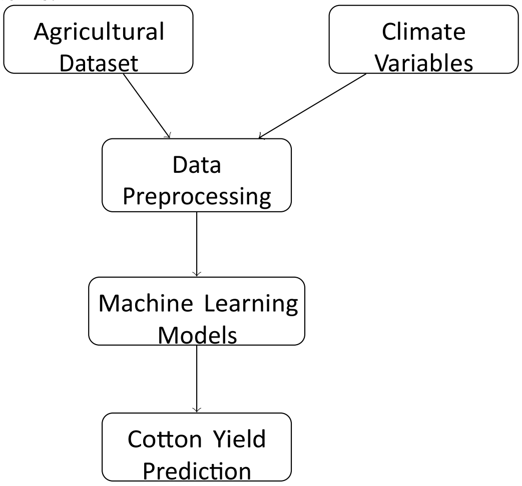

Fig. 1. Conceptual framework for machine learning-based cotton yield prediction.

Overall, the integration of machine learning techniques with agricultural datasets offers a promising solution for improving crop yield prediction accuracy. Such models can assist farmers and policymakers in understanding yield trends, mitigating climate risks, and implementing sustainable agricultural practices.

II. Literature Review

Several research works have focused on the interplay between climate variability, machine learning methodologies, and crop yield prediction. The increasing trend of agricultural and environmental data has enabled scientists to apply complex machine learning algorithms to understand crop yield patterns and provide insights for decision-making in the current agricultural sector.

Zhang et al. [12] used boosted tree regression and artificial neural network algorithms to predict the yield of upland rice under climate change scenarios in the Sahel region of Africa. The results of the research indicate the potential of machine learning algorithms to detect non-linear relationships between climatic factors and crop yield patterns. In a similar direction, Kumar and Singh [13] emphasized the role of microbial diversity and microbial biostimulants in enhancing agricultural sustainability and crop yield under changing climatic conditions. Similarly, the impact of climate change has also been extensively studied for its effects on agricultural productivity, directly or indirectly. In a research study, Erda et al. [14] examined the impact of climate change and CO_2 fertilization on crop yield and quality in China, which concluded that increased CO_2 levels can stimulate photosynthesis and crop yield to some extent. In a similar research study, Lawlor and Mitchell [15] concluded that increased CO_2 levels play a crucial role in the rate of photosynthesis of crop plants, thereby enhancing agricultural productivity.

In another research study, Chandio et al. [16] examined the short- and long-term effects of climate change on agricultural productivity using empirical data from China, which concluded that temperature variability and precipitation patterns play a crucial role in crop yield patterns. Similarly, Rossi et al. [17] emphasized the complexities involved in the decisionmaking process in the current agricultural sector and the role of technology in crop protection and yield enhancement. In recent times, precision agriculture technologies have become a prominent field of study. Abioye et al. [18] have shown the application of machine learning algorithm-based precision irrigation management and digital farming systems. It has been demonstrated that ML-driven precision irrigation management may greatly increase crop productivity and water economy. Sharma et al. [19] have conducted a thorough review of the application of machine learning in precision agriculture. They have shown that precision agriculture can be effectively managed using ML algorithms.

Machine learning has gained increased attention in other scientific disciplines as well. This increased attention is due to its capabilities in analyzing complex data sets. Jin et al. [20] have conducted a review of recent developments in the application of machine learning algorithms in experimental mechanics. They have shown the versatility of ML in solving complex engineering and environmental science problems. In the field of agriculture, Van Klompenburg et al. [21] have conducted a systematic literature review on crop yield prediction using ML. They have shown that ML models are significantly better than traditional statistical models in solving several agricultural prediction tasks.

Machine learning has been combined with remote sensing and spatial analysis techniques to improve the application of these techniques in agriculture. Tariq et al. [22] have combined optical remote sensing with decision tree and random forest algorithm-based spatial analysis to study the patterns of croplands. Similarly, Yang et al. [23] have proposed an optimal sampling algorithm to improve digital soil mapping using ML predictive uncertainty. Recent advances in artificial intelligence have also increased the rate of adoption of machine learning approaches in the agricultural domain. Shaikh et al. [24] presented the role of artificial intelligence and machine learning in smart farming and precision agriculture. Morales and Villalobos [25] emphasized the potential of machine learning

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 10-16, September 2026. ISSN (Online): 2456-7361

approaches for crop yield prediction using historical agricultural datasets.

Similarly, large-scale agricultural forecasting models have also been developed using machine learning techniques. In this context, Paudel et al. [26] have demonstrated the application of ML algorithms in the development of large-scale crop yield forecasting models using climatic and agricultural data. In the case of cotton crops, Yildirim et al. [27] have applied artificial neural networks in the development of crop yield prediction models in regions with limited agricultural data.vIn this context, other machine learning techniques have also been explored for crop yield prediction models. In this regard, Fashoto et al. [28] have applied multiple linear regression techniques in the development of maize yield prediction models. Similarly, Reddy and Kumar [29] have applied ML algorithms in improving the accuracy of crop yield forecasting models. Falamarzi et al. [30] have applied artificial and wavelet neural networks in the development of models for the estimation of evapotranspiration using temperature and wind speed data.

In this context, aerial imagery techniques have also been applied in the development of crop yield prediction models. In this regard, Feng et al. [31] have applied UAV-based multisensor imagery techniques in the development of crop yield prediction models for cotton crops. This demonstrates the application of remote sensing technologies in the development of crop yield prediction models. In this context, Li et al. [32] have applied ensemble learning techniques in the development of crop yield prediction models for soybean crops using meteorological data.

From the above literature, it is evident that there is a strong focus on the application of machine learning, remote sensing, and climatic data analysis in the development of crop yield prediction models. This demonstrates the application of data-driven models in improving the accuracy of crop yield prediction models.

III. METHODOLOGY

This section presents the methodology used to develop the cotton yield prediction framework. The proposed approach integrates historical agricultural production data, climatic variables, and machine learning algorithms to estimate cotton production patterns. The methodology consists of dataset collection, preprocessing, feature engineering, model training, and performance evaluation.

A. Dataset Description

The dataset used in this study consists of historical agricultural and climatic records related to cotton production. The dataset contains observations collected from different districts across agricultural regions where cotton cultivation is a major economic activity. Each record in the dataset represents cotton production for a specific district, season, and year.

The dataset spans multiple years ranging from 2010 to 2025 and includes several environmental and production related variables. These variables capture important climatic and agricultural factors that influence crop productivity. The dataset contains information related to cultivated area, annual rainfall, temperature patterns, humidity, fertilizer usage, and cotton production quantities.

TABLE I. Summary of Previous Studies on Machine Learning for Crop Yield Prediction

Ref.	Crop/Area	Method Used	Key Contribution
[12]	Rice / Sa-hel	Boosted Tree, ANN	Forecasted crop yield under climate change conditions
[14]	China Crops	Climate Modeling	Studied CO_2 fertilization effects on crop yield
[18]	Precision Agriculture	ML Irrigation Models	Improved irrigation efficiency using ML
[21]	Global Agriculture	ML Review	Systematic review of crop yield prediction methods
[26]	Largescale Agriculture	ML Forecasting Models	Developed large-scale crop yield prediction framework
[27]	Cotton	Artificial Neur al Networks	Predicted cotton yield in data-scarce regions
[31]	Cotton	UAV Imaging	Cotton yield estimation using aerial imagery
[32]	Soybean / China	Ensemble Learning	Improved crop prediction using ensemble ML models

The main attributes used in the dataset include:

- Crop Year
- District
- Season
- Month
- Area under cultivation
- Production
- Yield
- Annual Rainfall
- Maximum Temperature
- Minimum Temperature
- Average Temperature
- Humidity
- Fertilizer usage

The reason why these factors are chosen is that they have a big impact on the final output and crop growth cycles.

B. Data Preprocessing

Before applying machine learning models, the dataset was preprocessed to ensure data consistency and quality. The preprocessing stage involved cleaning the dataset, handling missing values, encoding categorical variables, and preparing features for model training.

Missing numerical values were replaced using the median of each respective feature, which helps maintain the distribution of the dataset while preventing bias caused by extreme values. Categorical attributes such as district and season were encoded into numerical representations using label encoding.

Mathematically, missing value replacement can be expressed as

$$x_i^* = \begin{cases} x_i, & \text{if value exists} \\ \tilde{x}, & \text{if value is missing} \end{cases} \quad (1)$$

where x_i represents the original feature value and $\tilde{x}$ repre-

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 10-16, September 2026. ISSN (Online): 2456-7361

sents the median value of that feature.

C. Feature Engineering

Feature engineering was performed to enhance the predictive capability of the models. Additional variables were derived from existing climatic attributes to better capture environmental variability affecting cotton production.

One derived feature used in the study is the temperature range, which captures daily thermal variation:

$$Trange = Tmax - Tmin \quad (2)$$

where T_{max} represents maximum temperature and T_{min} represents minimum temperature.

The input feature vector used for model training can be represented as

$$X = \{x_1, x_2, x_3, \ldots, x_n\} \quad (3)$$

where X represents the set of climatic and agricultural input variables.

The target variable for prediction is cotton production, defined as

$$Y = f(X) \quad (4)$$

where $f(.)$ represents the mapping function learned by the machine learning models.

D. Machine Learning Models

Three machine learning algorithms were used to develop the cotton yield prediction models:

- Random Forest Regression
- Gradient Boosting Regression
- Extreme Gradient Boosting (XGBoost)

These algorithms were selected because of their ability to model nonlinear relationships and handle complex datasets.

1) *Random Forest Regression:* Random Forest is an ensemble learning technique that combines multiple decision trees to improve prediction accuracy. The prediction of a Random Forest model is obtained by averaging the predictions from individual trees.

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \quad (5)$$

where $T_i(x)$ represents the prediction from the i^{th} decision tree and N is the total number of trees.

2) *Gradient Boosting:* Gradient Boosting builds models sequentially by minimizing prediction errors using gradient descent. Each subsequent model corrects the errors of the previous model.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (6)$$

where $h_m(x)$ represents the weak learner added at iteration m and γ_m represents the learning rate.

3) *Extreme Gradient Boosting (XGBoost):* XGBoost is an optimized implementation of gradient boosting designed for high computational efficiency. The objective function minimized by XGBoost is

$$L(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k} \Omega(f_k) \quad (7)$$

where l represents the loss function and Ω represents the regularization term used to control model complexity.

E. Training Procedure

The dataset was divided into training and testing subsets to evaluate the performance of the machine learning models. Approximately 85% of the dataset was used for training while the remaining 15% was used for testing.

The training process follows the steps outlined in Algorithm III-E.

Machine Learning Training Procedure

Load cotton production dataset Clean dataset and remove inconsistencies Handle missing values Encode categorical variables Generate engineered features Split dataset into training and testing sets each machine learning model Train model on training dataset Predict cotton production on testing dataset Evaluate model performance Compare results of all models

F. Model Evaluation

The performance of the models was evaluated using three standard regression metrics.

The coefficient of determination (R^2) measures the proportion of variance explained by the model:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (8)$$

Mean Absolute Error (MAE) is defined as

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (9)$$

Root Mean Square Error (RMSE) is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (10)$$

These evaluation metrics provide insights into prediction accuracy and model reliability.

G. Proposed Framework

The overall workflow of the proposed cotton production prediction system is illustrated in Fig. 2. The framework integrates agricultural data preprocessing with machine learning training to generate accurate production forecasts.

The framework integrates agricultural datasets with machine learning techniques to develop accurate predictive models capable of supporting data-driven decision-making in modern agriculture.

IV. Results and Discussion

In this section, the experimental results obtained based on the machine learning models used for predicting cotton production will be presented. In this study, three different machine learning models were used for predicting cotton production, and they are Random Forest, Gradient Boosting, and XGBoost. The performances of these models were evaluated based on different regression metrics such as coefficient of determination (R^2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

A. Model Performance Comparison

Table II presents the overall performance comparison of the three models.

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 10-16, September 2026. ISSN (Online): 2456-7361

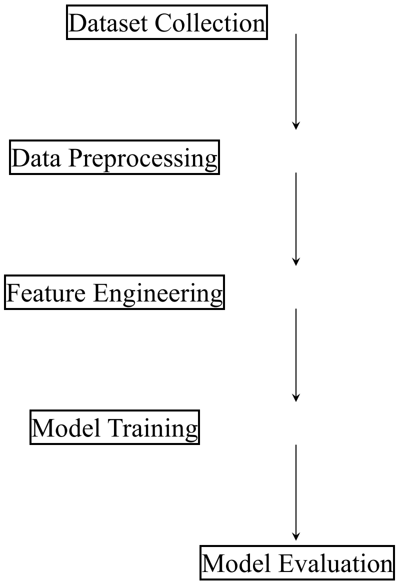

Fig. 2. Workflow of the proposed cotton yield prediction methodology

TABLE II. Performance Comparison of Machine Learning Models

Model	R^2 Score	MAE	RMSE
Random Forest	0.8867	33.16	44.25
Gradient Boosting	0.9341	24.82	33.74
XGBoost	0.9664	17.56	24.10

From the results presented in Table II, it can be observed that the XGBoost model achieved the highest prediction accuracy with an R^2 score of 0.9664. Gradient Boosting also performed well with an R^2 score of 0.9341, while Random Forest showed comparatively lower performance with an R^2 score of 0.8867. Similarly, the MAE and RMSE values indicate that XGBoost produced the lowest prediction error among the three models.

B. Random Forest Model Results

The prediction performance of the Random Forest model is illustrated in Fig. 3. The scatter plot compares the actual cotton production values with the predicted values generated by the Random Forest model.

From Fig. 3, it can be observed that most of the predicted values follow the general trend of the actual production values. However, some variations are visible for certain samples, which leads to higher prediction errors compared to the other models.

The feature importance obtained from the Random Forest model is shown in Fig. 4.

The feature importance graph indicates that the cultivated area is the most influential variable affecting cotton production. Crop year, district information, and fertilizer usage also contribute significantly to the prediction process.

C. Gradient Boosting Model Results

The performance of the Gradient Boosting model is shown in Fig. 5.

The predicted values in Fig. 5 show a stronger alignment with the actual production values compared to the Random Forest model. This indicates that Gradient Boosting is better able to capture the relationship between climatic variables and cotton production.

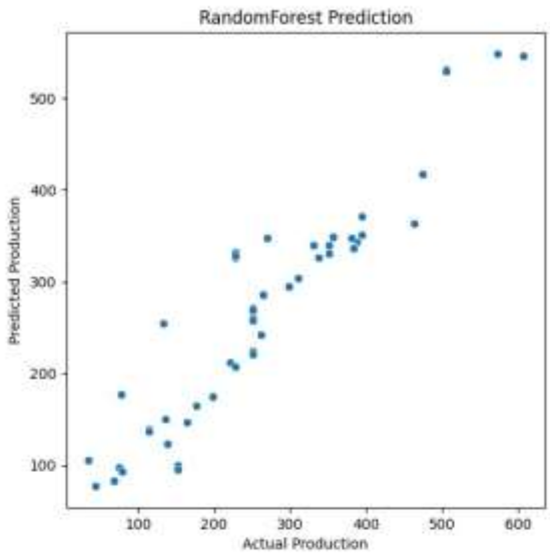

Fig. 3. Random Forest prediction results

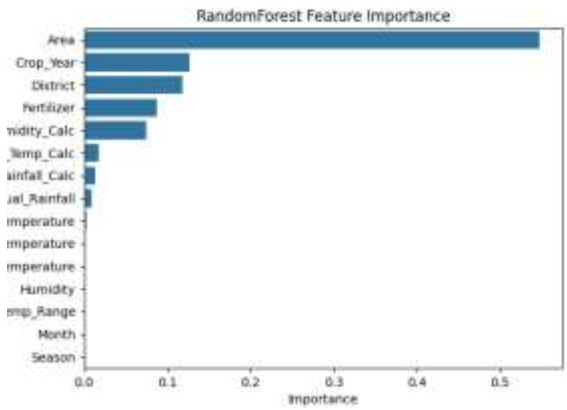

Fig. 4. Random Forest feature importance

The feature importance obtained from the Gradient Boosting model is presented in Fig. 6.

The results show that the cultivated area and crop year are the most dominant variables influencing cotton yield prediction. Other factors such as humidity, district information, and fertilizer usage also contribute to the model's prediction capability.

D. XGBoost Model Results

The prediction performance of the XGBoost model is illustrated in Fig. 7.

The results indicate that the XGBoost model produces predictions that closely match the actual cotton production values. Most of the points are distributed near the diagonal line, which demonstrates strong agreement between actual and predicted values.

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 10-16, September 2026. ISSN (Online): 2456-7361

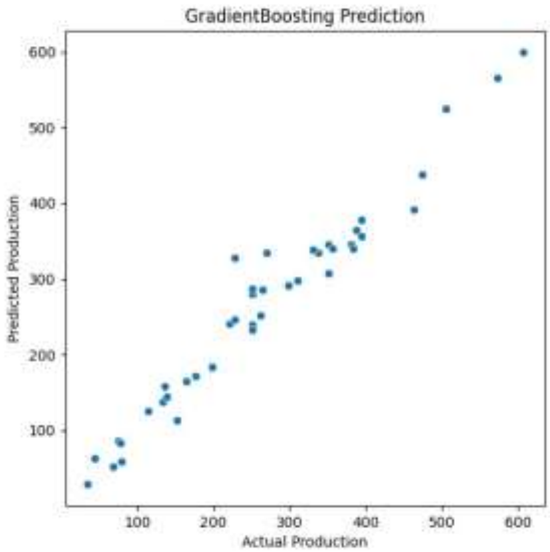

Fig. 5. Gradient Boosting prediction results

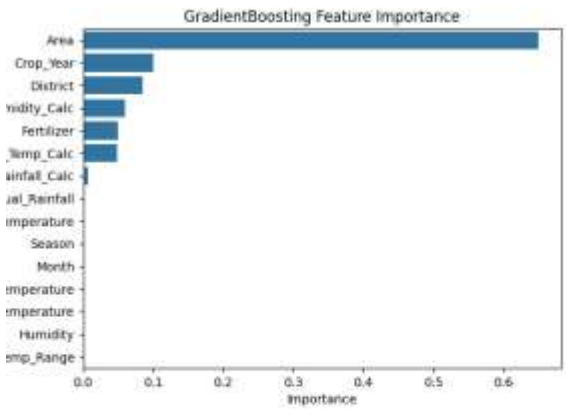

Fig. 6. Gradient Boosting feature importance

The feature importance analysis for the XGBoost model is shown in Fig. 8.

Similar to the previous models, the cultivated area is identified as the most important factor influencing cotton production. Fertilizer usage and crop year also show strong importance in determining production levels. Climatic variables such as rainfall, humidity, and temperature contribute moderately to the prediction process.

E. Discussion

The results of the study show that machine learning algorithms can be used to predict the production of cotton by considering climatic and agricultural factors. Among the three algorithms used in the study, the XGBoost algorithm has the highest accuracy in predicting the production of cotton, followed by the Gradient Boosting algorithm and the Random Forest algorithm.

The results of the feature importance show that the cultivated area is the most influential factor in the production of cotton. This is expected, as the output of the crop can be directly correlated with the cultivated area. Other factors, like the use of fertilizer, crop year, and district, also play a role in the outcome of the crop. The results of the study show that machine learning algorithms can be used to understand the complex relationship between climatic factors and the production of cotton, and hence can be a useful tool in the planning of the crop.

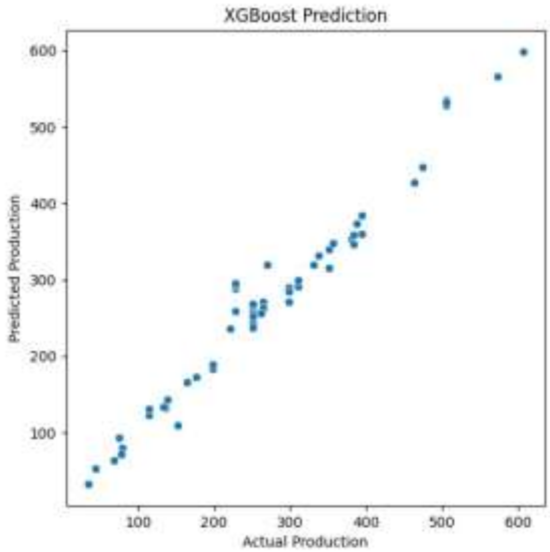

Fig. 7. XGBoost prediction results

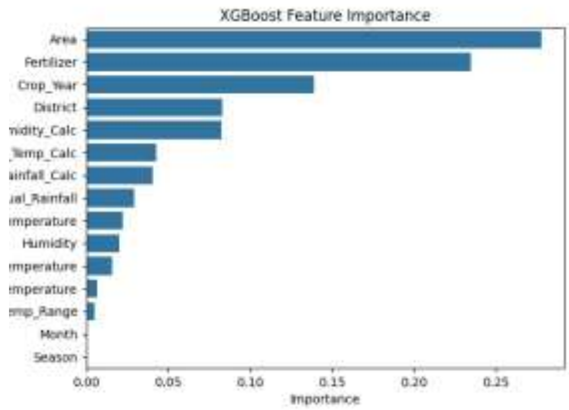

Fig. 8. XGBoost feature importance

V. Conclusion

In this paper, a machine learning-based framework for predicting cotton production is proposed based on historical agricultural and climatic data. The dataset includes a number of environmental and agricultural-related parameters, including rainfall, temperature, humidity, fertilizer usage, and cultivated area, among others. The dataset was preprocessed and feature engineered to improve the quality of the data for the machine learning model. The machine learning models, including Random Forest, Gradient Boosting, and XGBoost, were implemented for the prediction of cotton production. The results of the experiment show that the XGBoost model performed better than the other machine learning models, with a high accuracy of 0.9664 R^2, while the Gradient Boosting and Random Forest models were also able to perform well in the prediction of cotton production. The feature importance of the machine learning model revealed that the cultivated area, fertilizer usage, and crop year were the most influential parameters for the prediction of cotton production. The results of the experiment show that machine learning models can perform well in predicting the complex relationship between climatic parameters and agricultural production, which can be helpful for farmers, policymakers, and agricultural planners to make informed decisions for improving crop production and promoting better agricultural management practices.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Dr. Lal Chand, Assistant Professor, for his valuable guidance, continuous support, and helpful suggestions throughout this research work. His expertise and encouragement greatly helped in completing this study successfully. The authors also thank the Department of Computer Science and the institute for providing the necessary resources and environment to carry out this research work on mammogram image classification using deep learning.

REFERENCES

[1] M. A. Ali, J. Farooq, A. Batool, A. Zahoor, F. Azeem, A. Mahmood, and K. Jabran, “Cotton production in Pakistan,” in *Cotton Production*, 2019, pp. 249–276.
[2] Textile Exchange, “Organic cotton market report,” Available: http://textileexchange.org, 2011, pp. 65–66.
[3] M. P. Bange, J. T. Baker, P. J. Bauer, K. J. Broughton, G. A. Constable, Q. Luo, D. M. Oosterhuis, Y. Osanai, P. Payton, and D. T. Tissue, “Climate change and cotton production in modern farming systems,” CABI, vol. 6, 2016.
[4] M. A. Naseem and Z. Shahid, “Enhancing cotton production in Pakistan: evaluating and addressing challenges in agribusinesses,” *Journal of Sustainable Food and Agribusiness*, vol. 1, no. 1, pp. 41–47, 2023.
[5] J. Baffes, “Markets for cotton by-products: Global trends and implications for African cotton producers,” *World Bank Policy Research Working Paper*, no. 5355, 2010.
[6] F. Shuli, A. H. Jarwar, X. Wang, L. Wang, and Q. Ma, “Overview of the cotton in Pakistan and its future prospects,” *Pakistan Journal of Agricultural Research*, vol. 31, no. 4, p. 396, 2018.
[7] Y. Kang, S. Khan, and X. Ma, “Climate change impacts on crop yield, crop water productivity and food security–a review,” *Progress in Natural Science*, vol. 19, no. 12, pp. 1665–1674, 2009.
[8] M. I. Ahmad and H. Ma, “Climate change and livelihood vulnerability in mixed crop–livestock areas: the case of province Punjab, Pakistan,” *Sustainability*, vol. 12, no. 2, p. 586, 2020.
[9] K. E. Trenberth, *Climate System Modeling*. Cambridge, U.K.: Cambridge University Press, 1992.
[10] J. Gornall, R. Betts, E. Burke, R. Clark, J. Camp, K. Willett, and A. Wiltshire, “Implications of climate change for agricultural productivity in the early twenty-first century,” *Philosophical Transactions of the Royal Society B*, vol. 365, no. 1554, pp. 2973–2989, 2010.
[11] S. Peng, C. Wang, Z. Li, K. Mihara, K. Kuramochi, Y. Toma, and R. Hatano, “Climate change multi-model projections in CMIP6 scenarios in central Hokkaido, Japan,” *Scientific Reports*, vol. 13, no. 1, p. 230, 2023.
[12] L. Zhang, S. Traore, J. Ge, Y. Li, S. Wang, G. Zhu, Y. Cui, and G. Fipps, “Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel,” *Computers and Electronics in Agriculture*, vol. 166, p. 105031, 2019.
[13] P. Kumar and R. P. Singh, “Microbial diversity and multifunctional microbial biostimulants for agricultural sustainability,” in *Climate Resilient Environmental Sustainable Approaches: Global Lessons and Local Challenges*, 2021, pp. 141–184.
[14] L. Erda, X. Wei, J. Hui, X. Yinlong, L. Yue, B. Liping, and X. Liyong, “Climate change impacts on crop yield and quality with CO_2 fertilization in China,” *Philosophical Transactions of the Royal Society B*, vol. 360, no. 1463, pp. 2149–2154, 2005.
[15] D. Lawlor and R. Mitchell, “The effects of increasing CO_2 on crop photosynthesis and productivity: A review of field studies,” *Plant, Cell and Environment*, vol. 14, no. 8, pp. 807–818, 1991.
[16] A. A. Chandio, Y. Jiang, A. Rehman, and A. Rauf, “Short and longrun impacts of climate change on agriculture: An empirical evidence from China,” *International Journal of Climate Change Strategies and Management*, vol. 12, no. 2, pp. 201–221, 2020.
[17] V. Rossi, T. Caffi, and F. Salinari, “Helping farmers face the increasing complexity of decision-making for crop protection,” *Phytopathologia Mediterranea*, pp. 457–479, 2012.
[18] E. A. Abioye, O. Hensel, T. J. Esau, O. Elijah, M. S. Z. Abidin, A. S. Ayobami, O. Yerima, and A. Nasirahmadi, “Precision irrigation management using machine learning and digital farming solutions,” *AgriEngineering*, vol. 4, no. 1, pp. 70–103, 2022.
[19] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, “Machine learning applications for precision agriculture: A comprehensive review,” *IEEE Access*, vol. 9, pp. 4843–4873, 2020.
[20] H. Jin, E. Zhang, and H. D. Espinosa, “Recent advances and applications of machine learning in experimental solid mechanics: A review,” *Applied Mechanics Reviews*, vol. 75, no. 6, p. 061001, 2023.
[21] T. Van Klompenburg, A. Kassahun, and C. Catal, “Crop yield prediction using machine learning: A systematic literature review,” *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.
[22] A. Tariq, J. Yan, A. S. Gagnon, M. Riaz Khan, and F. Mumtaz, “Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest,” *Geo-Spatial Information Science*, vol. 26, no. 3, pp. 302–320, 2023.
[23] H. Yang, H. Lim, H. Moon, Q. Li, S. Nam, J. Kim, and H. T. Choi, “Simple optimal sampling algorithm to strengthen digital soil mapping using the spatial distribution of machine learning predictive uncertainty: A case study for field capacity prediction,” *Land*, vol. 11, no. 11, p. 2098, 2022.
[24] T. A. Shaikh, T. Rasool, and F. R. Lone, “Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming,” *Computers and Electronics in Agriculture*, vol. 198, p. 107119, 2022.
[25] A. Morales and F. J. Villalobos, “Using machine learning for crop yield prediction in the past or the future,” *Frontiers in Plant Science*, vol. 14, p. 1128388, 2023.
[26] D. Paudel, H. Boogaard, A. de Wit, S. Janssen, S. Osinga, C. Pylianidis, and I. N. Athanasiadis, “Machine learning for large-scale crop yield forecasting,” *Agricultural Systems*, vol. 187, p. 103016, 2021.
[27] T. Yildirim, D. N. Moriasi, P. J. Starks, and D. Chakraborty, “Using artificial neural network (ANN) for short-range prediction of cotton yield in data-scarce regions,” *Agronomy*, vol. 12, no. 4, p. 828, 2022.
[28] S. G. Fashoto, E. Mbunge, G. Ogunleye, and J. V. den Burg, “Implementation of machine learning for predicting maize crop yields using multiple linear regression and backward elimination,” *Malaysian Journal of Computing*, vol. 6, no. 1, pp. 679–697, 2021.
[29] D. J. Reddy and M. R. Kumar, “Crop yield prediction using machine learning algorithm,” in *Proc. 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, 2021, pp. 1466–1470.
[30] Y. Falamarzi, N. Palizdan, Y. F. Huang, and T. S. Lee, “Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks,” *Agricultural Water Management*, vol. 140, pp. 26–36, 2014.
[31] A. Feng, J. Zhou, E. D. Vories, K. A. Sudduth, and M. Zhang, “Yield estimation in cotton using UAV-based multi-sensor imagery,” *Biosystems Engineering*, vol. 193, pp. 101–114, 2020.
[32] Q.-C. Li, S.-W. Xu, J.-Y. Zhuang, J.-J. Liu, Z. Yi, and Z.-X. Zhang, “Ensemble learning prediction of soybean yields in China based on meteorological data,” *Journal of Integrative Agriculture*, vol. 22, no. 6, pp. 1909–1927, 2023.