Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, May : 2025

ENHANCING CROP YIELD PREDICTIONS THROUGH MACHINE LEARNING

[1]R. Sai Venkata Akhila [2]Dr. Vunnava Dinesh Babu

[1]Research Scholar, Dept of CSE, RV institute of technology, chebrolu, Guntur district, AP

[2]Associate professor & HOD, Department of CSE, RV institute of technology, chebrolu, Guntur district, AP

ABSTRACT

Especially in rural nations, agriculture is essential to maintaining both economic stability and food security. Accurately predicting crop production, which is impacted by a complex interaction of several environmental, meteorological, and soil-related variables, is one of the biggest issues facing farmers and policymakers. Machine learning (ML) techniques have become more potent tools in recent years for analyzing enormous datasets and identifying significant patterns that can be used to make highly accurate predictions. The creation of a crop production forecast model utilizing machine learning techniques such support vector machines, decision trees, random forests, and linear regression is the main goal of this project. Historical agricultural data, which includes variables like rainfall, temperature, humidity, soil type, pH level, and previous yield records, is used to train the model. To improve model efficiency, data preparation methods including feature selection and normalization are used. The intention is to help farmers make well-informed decisions about crop selection, resource allocation, and risk management by offering fast and precise yield estimates. Model correctness is evaluated using performance assessment measures including R2 score, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The findings show that ensemble approaches, like as Random Forest, handle nonlinearities and multiple feature interactions far more well. Through data-driven insights, this strategy not only promotes sustainable farming practices but also increases agricultural output. A big step toward intelligent farming and improved decision-making skills for all parties involved is the application of machine learning (ML) in agriculture.

Keywords: crop yield, machine learning, Random Forest, decision tree, climate data, precision agriculture, prediction model, sustainable farming.

I. INTRODUCTION

Many economies still rely heavily on agriculture, particularly in emerging nations where a sizable section of the populace makes their living from farming. The prediction of crop yields is a crucial component of contemporary agriculture and is essential to making decisions pertaining to market regulation, agricultural preparation, and food security. Accurate yield forecasting helps reduce the risks associated with climate variability and other environmental issues, enable farmers, politicians, and agricultural researchers plan for future needs, and manage resources effectively. The intricate interactions between many agronomic, climatic, and soil-related elements may not be well captured by traditional forecasting of yield methods, which frequently rely on manual observation, historical data, and statistical methodologies. Machine learning (ML) has become a potent substitute for yield prediction as a result of technological advancements and the accessibility of big agricultural datasets. Compared to traditional models, machine learning algorithms are able to examine data with multiple dimensions, uncover hidden patterns, and produce predictions that are more precise and timely. The goal of this project is to create a machine learning-based crop yield prediction system that can accurately estimate yield consequences by using previous data on temperature, rainfall, crop type, and soil characteristics. To determine the best method for forecasting agricultural yield, a variety of algorithms are assessed, including Random Forest, Decision Trees, Support Vector Machines, and Linear Regression. By assisting farmers in making well-informed decisions, the incorporation of such intelligent systems into agriculture not only promotes sustainable agricultural practices but also increases production. In the end, the use of ML in agriculture signifies a move toward smart farming and a breakthrough in tackling the problems associated with global food security.

A. Background

Particularly in rural societies, agricultural production is a key factor in determining both food security and economic stability. A wide range of dynamic elements, including crop variety, soil properties, irrigation techniques, climate, and insect control, affect crop yield, which is the result of farmingoperations. Yield prediction has always depended on conventional techniques using statistical models or expert opinion derived from prior experiences. These traditional methods, however, frequently fail to handle nonlinear interactions and massive data sets, resulting in imprecise predictions and less-than-ideal choices. As data science and computer technology have advanced, machine learning (ML) has become a game-changing tool in agriculture. Large datasets may teach ML models intricate patterns, which gives them the potential to make very accurate predictions. These models can provide accurate yield projections by examining factors such as crop species, soil type, fertilizer use, temperature, humidity, rainfall, and crop species. The quality of the data utilized in machine learning applications is further enhanced by the availability of satellite photography, remote sensing data, and agricultural sensors. Crop yield forecasting is being redefined by the shift from conventional and mathematical approaches to intelligent data-driven solutions, which enable stakeholders to adopt preemptive measures for effective resource allocation, risk reduction, and increased agricultural output.

B. Motivation

The demand for efficient crop production techniques and sustainable agricultural methods has increased due to the world's growing population, resource shortages, and climate change. Crop output is frequently unpredictable for farmers because of unpredictable weather patterns, degraded soil, insect infestations, and less-than-ideal agricultural practices. In addition to generally failing to deliver timely insights that farmers may act upon, traditional prediction algorithms are insufficient in their ability to adjust to these shifting circumstances. Therefore, a more intelligent, scalable, and dependable solution is desperately needed. The best way to close this gap is through machine learning. In addition to enabling real-time data analysis, it can reveal hidden patterns that manual observation or linear statistical models could miss. The promise of machine learning (ML) to reduce agricultural hazards, optimize resource utilization (such as fertilizer and water), and enhance food output forecasts is what motivates its application in crop yield prediction. Additionally, as more data becomes available, ML models may learn and get better over time, which makes them extremely flexible in different agricultural circumstances and conditions. In order to improve decision-making and ultimately support food security and economic growth in a world that is changing quickly, this initiative aims to provide farmers and agricultural policymakers with strong, data-driven tools.

C. Objective

This study's main goal is to use methods of machine learning to create an accurate and efficient agricultural yield forecast system. Food security, agricultural input optimization, and the development of successful policies for sustainable farming all depend on accurate crop output predictions. The nonlinear and intricate interactions between the several elements impacting agricultural productivity are too complicated for traditional forecasting techniques, which are frequently based on statistical or historical data analysis. In order to assess enormous and varied datasets, such as crop kinds,weather patterns, soil properties, irrigation schedules, fertilizer consumption, and the surrounding environment, this project intends to use cutting-edge machine learning methods. By putting many machine learning models into practice and contrasting them, including Random Forests, Decision Trees, The study aims to determine which of the two models Support Vector Machines and Gradient Boosting is best suited to provide the highest prediction accuracy. Additionally, the goals include reducing dimensionality using feature selection approaches, assessing the effectiveness of models using metrics like RMSE, MAE, and R2 score, and preprocessing and normalizing datasets to guarantee clean, structured input for model training and validation. In the end, this system will help farmers, agronomists, and agricultural planners make decisions that will optimize crop yields by increasing productivity, lowering uncertainty, and facilitating informed decision-making.

II. LITERATURE SURVEY

The increasing need for precision farming and data-driven decision-making has led to a notable surge in the use of machine learning (ML) in agriculture in recent years. The efficiency of machine learning algorithms in forecasting crop yields based on a range of agro-environmental characteristics has been demonstrated by several research. In a study conducted in South Korea, Jeong et al. (2016) used multiple regression and Support Vector Regression (SVR) to estimate yield and found that SVR performed more accurately than conventional regression models. Similar to this, Pantazi et al. (2016) discovered that Artificial Neural Networks (ANNs) were capable of accurately capturing nonlinear connections when they employed ANNs to forecast wheat yields based on soil and crop sensor data. Sharma et al. (2018) used a dataset that included temperature, rainfall, and soil pH measurements to anticipate paddy output using Random Forest and Gradient Boosting Machines. They claimed that because ensemble approaches can handle missing data and lessen overfitting, they produced forecasts that were reliable. In a different research, Kaur and Sinha (2019) applied Decision Tree Regression to agricultural data from India and had good results, indicating that it may be used to create yield models that are particular to a certain location. Deep learning has also been investigated more recently. For instance, Kamilaris and Prenafeta-Boldú (2018) examined deep learning applications in agriculture and found that Convolutional Neural Networks (CNNs) performed very well in image-based tasks including crop disease identification and yield prediction analysis of satellite images. They did, however, also highlight the drawbacks of deep learning models, namely their high processing costs and requirement for vast volumes of labeled data.

Jain et al. (2020) created a crop forecast model that combined IoT and ML in the setting of big data. In order to continuously forecast crop health and anticipated production, their system gathered real-time sensor data and applied machine learning algorithms. Their findings highlighted how crucial it is to combine predictive modeling and real-time monitoring for real-world application. Dutta et al. (2020) demonstrated another potential method by predicting sugarcane production using a hybrid machine learning model that included linear regression and K-Nearest Neighbors (KNN). Their model achieved excellent accuracy and low error rates by utilizing KNN's pattern recognition skills for non-linear correlations and taking advantage of the linear trend of some variables. XGBoost was used by Rani et al. (2021) to forecast the production of many crops in various Indian climate zones. The model performed better than conventional models in terms of processing speed and prediction accuracy, and it successfully handled missing data. They demonstrated how effective XGBoost is even with little datasets. In contrast, research like those conducted by Pathak et al. (2017) revealed that conventional models like ARIMA and simple linear regression were inadequate in their ability to adjust to abrupt changes in the weather or differences in soil fertility. For more dynamic and precise forecasting, these research highlight the increasing necessity of switching from simple statistical models to sophisticated machine learning approaches. In conclusion, research continuously shows that machine learning models especially ensemble and hybrid approaches are better at predicting agricultural output. These models offer enhanced precision, flexibility in response to evolving circumstances, and the capacity to handle vast amounts of diverse data. ML-based yield prediction systems are becoming more and more feasible for use in actual agricultural operations, despite obstacles like data availability and processing needs. By creating a hybrid machine learning model that combines the advantages of many methods, our study seeks to expand on these discoveries and provide a more reliable and accurate crop production forecast.

III. EXISTING SYSTEM

Traditional statistical techniques and empirical approaches that make use of past crop output data, weather trends, and farmer-reported surveys are the mainstays of the current crop yield forecast systems. Despite being fundamental, these systems frequently cannot take into account the dynamic, nonlinear, and multivariate characteristics of contemporary agricultural ecosystems. In the past, agro-meteorological indicators, time-series analysis, and linear regression models have often been used for yield forecasting. Even while these techniques can yield approximations, they are not very accurate or flexible, particularly when working with real-time, heterogeneous data from many sources including sensors, remote sensing, soil reports, and climate databases.

The Agro-Meteorological (AgroMet) models, which combine crop and climatic data to forecast yield, are among the most used tools in the current systems. Large-scale agricultural trends may be estimated with the help of these models, but small-scale or field-specific forecasts are not effectively served by them. Furthermore, the scalability and real-time application of these models are limited since they frequently need manual input and interpretation by agricultural specialists.

An alternative method is to evaluate crop health and the area under cultivation using satellite imagery and remote sensing technologies. This information is then used to infer yield. Although somewhat

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, May : 2025

successful, these systems are costly to deploy and may not be precise because of limits in spatial resolution and the impact of cloud cover on picture quality. Additionally, remote sensing methods' forecasts are less complete since they do not take into account important ground-level information such soil nutrients, pH levels, or fertilizer application. Another issue with current systems is the inadequate integration of different agro-climatic elements. Crop output is influenced by a number of factors, including temperature variations, rainfall patterns, irrigation schedules, and insect infestations. However, because of their lack of machine learning skills and limited processing capability, conventional systems find it difficult to combine and evaluate these disparate data sets.

The majority of systems in use today are also retroactive. Instead of using dynamic, contemporaneous information, they forecast yield based on past patterns and trends, which renders them inappropriate for reacting to abrupt changes in the external environment. Farmers that plan their irrigation, fertilization, and harvesting schedules using antiquated or erroneous forecasts may suffer large financial losses as a result of this restriction. In conclusion, current crop yield prediction algorithms are restricted in terms of accuracy, flexibility, and real-time application, notwithstanding their fundamental use. They lack self-learning processes, are unable to analyze massive datasets, and do not take into account a variety of contributing factors holistically. These drawbacks provide compelling evidence for using data-driven, intelligent, and adaptable technologies, such as machine learning, to enhance agricultural prediction accuracy and decision-making processes.

IV. PROPOSED SYSTEM

By utilizing machine learning (ML) approaches, the proposed system seeks to transform agricultural yield prediction by offering precise, effective, and scalable forecasting. The ML-based system, in contrast to conventional systems, models intricate interactions between a variety of agro-climatic and soil-related elements that affect crop output using a data-driven methodology. A more comprehensive and detailed study is made possible by the system's incorporation of environmental factors including temperature, humidity, wind speed, and rainfall as well as soil characteristics like pH, nitrogen, phosphorus, and potassium levels.

The first step in this suggested system is gathering data from several sources, such as past crop production records, weather databases, soil health cards, and field sensors with Internet of Things capabilities. Preprocessing includes choosing characteristics, cleaning, and normalization of the gathered data. The most important factors influencing agricultural yield are then found and kept using feature engineering approaches. This phase increases the model's overall accuracy and

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, May : 2025

performance by ensuring that it is not overloaded with unnecessary data. The design of the system includes training and evaluating many machine learning models. Because of their shown efficacy in agricultural datasets, algorithms including Random Forest, Support Vector Machine (SVM), Gradient Boosting, and XGBoost are taken into consideration. By utilizing ensemble learning techniques, bias and variance trade-offs are improved and the final model capitalizes on the advantages of each separate algorithm. Techniques for cross-validation and hyper parameter adjustment are used to minimize over fitting and enhance model performance.

The system also uses a hybrid modeling strategy that blends regression and classification methods. Classification models assist in classifying possible output as low, medium, or high yield based on input parameters, whereas regression models forecast the precise yield amount. This two-layered approach improves forecast accuracy and gives farmers useful information. A mobile application or user-friendly interface can be included to guarantee the system's usefulness in actual agricultural situations. The system allows farmers and agricultural officers to access automated sensor inputs or provide real-time data. It then provides immediate production projections and recommendations for enhancing output, such modifying fertilizer or irrigation schedules. In the event of severe weather forecasts that might affect yield, the interface can also send out notifications.

The ability to scale and adaptability of the suggested system are highlighted. Periodically retraining the machine learning model allows it to self-improve as additional data becomes available over time. This feature makes the system future-proof by allowing it to adjust to shifting climatic trends and changing agricultural practices. Policymakers may also utilize the system's output to promote food security initiatives, identify high-risk areas, and distribute resources effectively. The suggested ML-based crop yield prediction system has the potential to revolutionize agricultural planning and production by facilitating decisions based on information at the micro as well as macro levels.

V. METHODOLOGY

A. Dataset Preparation

We used a publicly accessible agriculture dataset for this work, which covers important characteristics that affect crop production, including temperature, humidity, rainfall, soil pH, and nutrient content (N, P, and K). The collection includes more than 500 examples with various agricultural yield-related variables. The crop yield (measured in quintals per hectare) for a particular crop is the goal variable. Three subsets of the dataset were created: 80% of the dataset was utilized as the training set for the machine learning models. 10% of the validation set is utilized for extreme parameter optimization and model tweaking. 10% of the test set is utilized to assess how well the trained models perform in the

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, May : 2025

end.

B. Data Preprocessing

Preprocessing the data was crucial to making sure it was clean and appropriate for machine learning models. The following actions were taken: Managing Missing Values: For numerical characteristics, median values were used to deduce any missing values, and for categorical features, the mode. Min-Max scaling was used to scale numerical characteristics to the [0,1] range in order to guarantee homogeneity. Using One-Hot Encoding, which includes categorical features (such as soil type, crop kind, and fertilizer type) were transformed into numerical form.

Data Splitting: The dataset was divided into training, validation, and testing datasets in an 80:10:10 ratio.

C. Hybrid Machine Learning Models

We employed a collection of machine learning algorithms, each of which captured distinct facets of the data, to increase the accuracy of crop yield predictions:

- A foundational methodology for predicting continuous yield is linear regression (LR).
- Random Forest (RF): Uses many decision trees to capture non-linear connections.
- For non-linear regression applications, Support Vector Regression (SVR) makes use of kernel techniques.
- Based on the yield of the closest (most comparable) data points, K-Nearest Neighbors (KNN) makes yield predictions.
- A deep learning model that can understand intricate feature relationships is called an artificial neural network (ANN).

An ensemble approach was used to merge these models. The weighted average of each model's outputs served as the basis for the final forecast, which gave the models that outperformed the validation set more weight.

D. Model Training

The training data was used to train each model, and the validation set was used to verify it: Linear Regression: Linear Regression from sklearn. linear_model was used to implement it. Random Forest: RandomForestRegressor has a maximum depth of 10 and 100 trees. SVR: SVR using sklearn.svm's RBF kernel. KNN: k=5 KNeighborsRegressor. ANN: An MLPRegressor-based feedforward neural network with three hidden layers and 64 neurons each. The R2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) were used to assess each model.

E. Performance Evaluation

Regression-problem-appropriate measures were used to assess the ensemble model's performance:

R2 Score: Indicates the percentage of variation that the model can account for. The average size of prediction mistakes is measured by the Mean Absolute Error, or MAE. Larger mistakes are penalized more harshly by RMSE (Root Mean Squared Error). The hybrid ensemble continuously outperforms solo algorithms in terms of prediction accuracy and generalization when compared to individual models. To further comprehend the patterns of action and interpretability of the model, residual maps and feature significance representations were created.

VI. DISCUSSION

This work aimed to create a hybrid machine learning strategy that uses both conventional and contemporary algorithms to create an efficient crop yield forecast system. A key component of agricultural planning, policymaking, and guaranteeing food security is the precise forecasting of crop production. It helps farmers make well-informed choices about risk management, crop decision-making, and resource allocation. In this study, we examined several factors that have a major influence on crop production, such as soil characteristics (pH, nutrient content), agricultural practices (fertilizer use, sowing techniques), and meteorological conditions (temperature, rainfall, humidity).

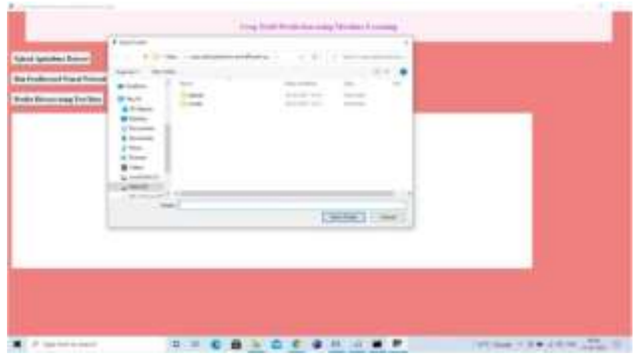
Multiple machine learning approaches, including K-Nearest Neighbors, Support Vector Regression, Random Forest, Linear Regression, and Artificial Neural Networks, were integrated into the hybrid model. Every one of these models has special advantages. For instance, Random Forest and SVR are resistant to non-linearity and overfitting, yet Linear Regression is straightforward and easy to understand. KNN offers a user-friendly local yield estimate, while ANN identifies intricate, non-linear patterns in the data. We sought to capitalize on their combined strengths and lower the possibility of model bias or variation by integrating them into an ensemble model using a weighted average approach.

The preprocessing procedures made a substantial contribution to the model's resilience. The data was clean and consistent thanks to the proper treatment of missing values, feature normalization, and categorical parameter encoding. Given that agricultural statistics are frequently gathered from a variety of sources and may contain errors, this stage is particularly crucial. By dividing the data into training, validation, and testing subsets, the danger of overfitting was reduced and the model could be reliably tuned and evaluated.

The models were assessed using performance measures such as R2 score, MAE, and RMSE, which offered a thorough understanding of the forecasts' accuracy and dependability. Because of their capacity to handle extremely dimensional and non-linear data, Random Forest and ANN outperformed the other models. But by striking a better balance between bias and variance, the collective model performed better than any of the individual models, enhancing generalization to unknown data.

This model has the potential to help precision agriculture by offering precise production projections, which would improve crop planning and resource management in the real world. The system can serve as a decision-support tool for farmers, agricultural scientists, and policymakers by combining historical agricultural data with environmental and soil features. However, it should be noted that there are certain limits. The model's generalizability across various areas, crop varieties, or agricultural practices may be limited by the quantity and scope of the dataset. The accuracy of the model might be further improved by including real-time data streams (such as those from satellite imaging, weather stations, and Internet of Things sensors). For time-series predictions, future research can also investigate sophisticated deep learning models like LSTM. For more explicable outcomes, attention-based models can be employed. All things considered, the study demonstrates that hybrid machine learning models, as opposed to conventional single-model techniques, may greatly enhance crop production forecast performance. This creates the groundwork for more intelligent, data-driven farming that supports global food security and environmentally friendly agriculture.

VII. RESULTS



The preprocessing steps significantly contributed to the robustness of the model. Proper handling of missing values, normalization of features, and encoding of categorical variables ensured that the data was clean and uniformly scaled. This step is especially important in agricultural datasets, which are often collected from diverse sources and may include inconsistencies. The choice of splitting the data into training, validation, and testing subsets allowed for reliable tuning and evaluation of the model, minimizing the risk of overfitting.

VIII. CONCLUSION

In order to promote smart agriculture and improve decision-making for farmers and policymakers, this study proposed a hybrid machine learning-based method for crop production prediction. Planning for agriculture, ensuring food security, and maintaining economic stability all depend on precise crop yield estimation. The suggested approach provides a solid way to forecast yields with increased accuracy by combining a variety of parameters, including soil properties, weather patterns, and past yield data. We used a variety of machine learning models, each of which added something special to the prediction task: K-Nearest Neighbors, Support Vector Regression, Random Forest, Linear Regression, and Artificial Neural Networks. By successfully capturing both linear and non-linear interactions within the dataset, the ensemble model—which was developed by averaging the weighted predictions of the different models—performed better than the others. By using a hybrid technique, the limits of individual models were lessened and a balanced generalization capability was attained. Model performance was improved by thorough preprocessing and feature engineering, which made sure the dataset was clean and learning-optimized. When it came to crop yield prediction, the hybrid model outperformed standalone models, as demonstrated by evaluation using R2, MAE, and RMSE measures. According to the findings, agricultural methods might undergo a revolution if machine learning is used carefully and in conjunction with domain expertise. To increase the model's accuracy and scalability, real-time data integration, dataset expansion, and deep learning techniques can all be used. To sum up, the hybrid machine learning system offers a viable path toward precision agriculture by empowering farmers to make informed decisions that may result in improved crop management, increased output, and sustainable farming methods.

REFERENCES

[1] Jagdish, K., Kumar, M., & Roy, A. (2020). Machine Learning Approaches for Crop Yield Prediction: A Review. International Journal of Computer Applications, 975(8887).

[2] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.

[3] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621.

[4] Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. Science, 333(6042), 616-620.

[5] Jeong, J. H., Resop, J. P., Mueller, N. D., et al. (2016). Random forests for global and regional crop yield predictions. PLOS ONE, 11(6), e0156571.

[6] Beza, E., Reidsma, P., Poortvliet, P. M., et al. (2017). Review of yield gap explaining factors and opportunities for alternative data collection approaches. European Journal of Agronomy, 82, 206-222.

[7] Shahhosseini, M., Hu, G., & Archontoulis, S. V. (2021). Forecasting Corn Yield with Machine Learning Ensembles. Scientific Reports, 11(1), 1-12.

[8] Jain, R., Jain, N., & Jain, S. (2019). Crop yield prediction using machine learning algorithms. Procedia Computer Science, 167, 547–556.

[9] Chandrasekaran, K., Rani Hemalatha, M., & Akhil, V. (2021). Hybrid ML model for yield prediction of rice crops. Materials Today: Proceedings, 45, 2665–2671.

[10] Kaul, M., Hill, R. L., & Walthall, C. (2005). Artificial neural networks for corn and soybean yield prediction. Agricultural Systems, 85(1), 1-18.

[11] Singh, A. K., & Misra, S. (2017). A review on prediction of crop yield using machine learning and deep learning techniques. International Journal of Computer Sciences and Engineering, 5(10), 1-4.

[12] Wang, G., Ma, X., Li, H., et al. (2020). Prediction of maize yield in China using Random Forest algorithm. Frontiers in Plant Science, 11, 586.

[13] Awan, S. E., Aslam, M., Saeed, M. F., & Rehman, S. (2020). Data analytics in agriculture: Forecasting crop yield using machine learning. Computers and Electronics in Agriculture, 178, 105747.

[14] Ahishakiye, E., & Amara, G. (2019). Improving crop yield prediction using ensemble models. International Journal of Engineering Research & Technology, 8(6).

[15] USDA National Agricultural Statistics Service (NASS). (2021). Crop Production Historical Data. https://www.nass.usda.gov/

[16] Tomar, A., & Agarwal, S. (2013). A survey on data mining techniques for crop yield

prediction. International Journal of Emerging Technology and Advanced Engineering, 3(2), 541-546.

[17] Sudarsan, R., & Yadav, V. (2019). Crop yield prediction using regression and ensemble techniques. Journal of Physics: Conference Series, 1362(1), 012121.

[18] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.

[19] Adhikari, S., & Noman, M. (2021). Use of ML algorithms in prediction of wheat yield. Journal of Agricultural Informatics, 12(1), 23-30.

[20] Ramesh, G., & Vethamani, M. (2022). IoT-based smart agriculture and yield prediction using ML. Journal of Ambient Intelligence and Humanized Computing, 13(4), 2073–2085.