IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ☵ Peer-reviewed & Refereed journal ☵ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.13397

Crop Yield Predication Using Extreme Machine Learning for Sustainable Agriculture

Trupti Baburao Bhoir[1], Neha Vinod Sankhe[2], Shifa Afsar Kureshi[3], Prerna Prakash Ahire[4] and Prof. Monika Samir Pathare[5]

Theem College of Engineering[1-5]

Abstract: Sustainable agriculture plays a vital role in ensuring global food security, environmental protection, and economic stability. Accurate crop yield prediction enables farmers and policymakers to make informed decisions regarding crop planning and resource management. This paper proposes a machine learning approach for crop yield prediction using the Extreme Learning Machine (ELM), a fast-neural network model with strong generalization capability. The dataset includes soil nutrients such as Nitrogen, Phosphorous, and Potassium along with environmental parameters like temperature, humidity, and soil type. Categorical variables are encoded and missing values are handled using median replacement. The ELM model uses a Single Layer Feedforward Network with randomly generated input weights and output weights calculated using the Moore–Penrose pseudoinverse. Performance is evaluated using RMSE and R^2 Score. A Streamlit-based interface is developed to allow users to input parameters and receive instant crop yield predictions.

Keywords: Crop Yield Prediction, Extreme Learning Machine (ELM), Machine Learning, Sustainable Agriculture, Soil Parameters, Climatic Factors, Precision Agriculture, Agricultural Technology, Streamlit, Real-time Prediction

I. INTRODUCTION

Agriculture plays a crucial role in the **global economy** and **food supply chain**. With the **increasing population** and growing **demand for food**, efficient **agricultural planning** is necessary. One of the key challenges in agriculture is **accurately predicting crop yield**.
Traditional crop yield prediction methods rely on **historical data** and **manual estimation**, which may not capture **complex relationships between soil nutrients and environmental conditions**.
Machine learning techniques can analyze **multiple agricultural parameters simultaneously** and provide **more accurate crop yield predictions**.
This study focuses on the **Extreme Learning Machine (ELM) algorithm** to predict **crop yield** based on **soil nutrients and climatic factors**.

II. LITERATURE REVIEW

Several researchers have explored machine learning techniques for improving crop yield prediction and agricultural decision-making. A comprehensive review on crop yield prediction highlights the use of machine learning models to analyze environmental conditions, soil nutrients, and crop growth parameters to estimate agricultural productivity.

Research on **wheat yield prediction** using machine learning combined with **climate and NDVI data** demonstrates that integrating satellite imagery with environmental data can significantly improve prediction accuracy. These approaches help farmers monitor crop growth and make better planning decisions.

Another study on **agro-technological systems in traditional agriculture** explains how modern technologies such as machine learning and data analytics can support farmers by providing insights into crop health, irrigation needs, and fertilizer usage.

Machine learning techniques have also been applied to **cotton yield prediction** using both field data and synthetic datasets. These studies show that predictive models can effectively estimate crop productivity under varying environmental conditions.
Furthermore, research on **crop recommendation and yield estimation systems** indicates that algorithms such as Random Forest, Support Vector Machines, and Neural Networks can assist farmers in selecting suitable crops and predicting expected yield.

DOI: 10.17148/IARJSET.2026.13397

Overall, previous studies demonstrate that **machine learning models provide accurate and efficient solutions for crop yield prediction**. However, many traditional neural network approaches require longer training times. Therefore, this research focuses on the **Extreme Learning Machine (ELM)** algorithm, which provides faster training and efficient prediction performance for agricultural datasets.

III. METHODOLOGY

The proposed system uses a machine learning approach based on the Extreme Learning Machine (ELM) algorithm to predict crop yield using soil and environmental parameters. The methodology consists of several stages including data collection, preprocessing, model training, and prediction**.**

A. Data Collection:

The dataset used in this research contains various soil and climatic parameters that influence crop productivity. The important features in the dataset include:

- Nitrogen (N)
- Phosphorous (P)
- Potassium (K)
- Temperature
- Humidity
- Moisture
- Soil Type
- Crop Type
- Fertilizer Name

The target variable in the dataset is crop yield measured in kilograms per hectare (kg/ha).

B. Data Preprocessing:

Before training the model, the dataset undergoes several preprocessing steps to improve data quality and ensure better model performance.

1. **Handling Missing Values:**
   Missing nutrient values are replaced using the **median value** of the respective column.
2. **Encoding Categorical Variables:**
   Features such as **soil type, crop type, and fertilizer name** are categorical in nature. These attributes are converted into numerical form using **Label Encoding**.
3. **Dataset Splitting:**
   The dataset is divided into **training and testing sets** to evaluate the performance of the prediction model.

C. Extreme Learning Machine (ELM) Model:

The Extreme Learning Machine (ELM) **is a** Single Layer Feedforward Neural Network (SLFN) designed for fast learning and efficient prediction. Unlike traditional neural networks, ELM randomly assigns the input weights and hidden layer biases**,** while the output weights are calculated analytically using the Moore–Penrose pseudoinverse**.**
This approach significantly reduces training time while maintaining good generalization performance.

D. Model Evaluation:

The performance of the model is evaluated using the following metrics:

- **Root Mean Squared Error (RMSE)** – measures prediction error.
- **R^2 Score (Coefficient of Determination)** – indicates how well the model explains the variance in crop yield.

These evaluation metrics help assess the accuracy and reliability of the crop yield prediction system.

IV. SYSTEM ARCHITECTURE

The proposed crop yield prediction system is designed to analyze soil nutrients and environmental parameters using a machine learning model. The system consists of several stages including data **input, preprocessing, model training, prediction, and result visualization**.
Initially, agricultural data containing soil nutrients and climatic parameters is collected and stored in a dataset. The dataset is then preprocessed to handle missing values and convert categorical variables into numerical format. After preprocessing, the cleaned data is used to train the **Extreme Learning Machine (ELM) model**.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13397

The trained model learns the relationship between **soil nutrients, environmental factors, and crop yield**. Once the model is trained, it can predict crop yield based on new input values provided by the user. A **Streamlit-based interface** allows farmers or users to enter soil and environmental parameters and obtain instant yield predictions.

System Workflow

1. Data Collection
2. Data Preprocessing
3. Feature Encoding
4. Model Training using ELM
5. Yield Prediction
6. Result Display through Streamlit Interface

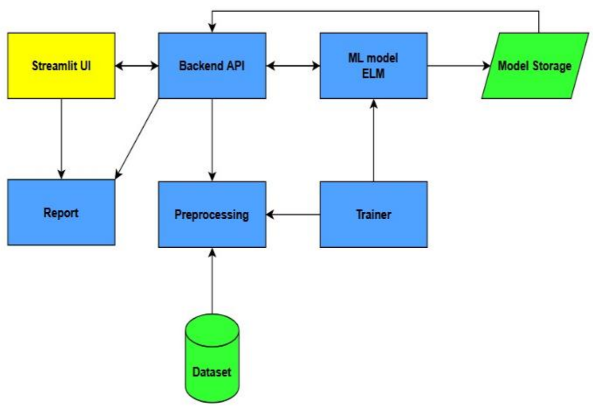

Fig. System Architecture

V. RESULTS AND DISCUSSION

The Extreme Learning Machine model was trained using the processed agricultural dataset containing soil nutrients and environmental factors. The model performance was evaluated using standard machine learning evaluation metrics.
The following metrics were used to evaluate the prediction performance:

- **Root Mean Squared Error (RMSE)** – Measures the average difference between predicted and actual crop yield values.
- **R^2 Score (Coefficient of Determination)** – Indicates how well the model explains the variance in crop yield data.

The results show that the Extreme Learning Machine model provides **fast training speed and reliable prediction accuracy** compared to traditional neural network models. The model is capable of analyzing multiple agricultural parameters simultaneously and generating yield predictions efficiently.

The integration of the **Streamlit interface** further improves usability by allowing users to input soil and environmental parameters and instantly receive predicted crop yield values.

VI. CONCLUSION

This research presents a machine learning-based crop yield prediction system using the Extreme Learning Machine (ELM) algorithm**.** The system analyzes soil nutrients and climatic parameters to estimate crop yield accurately.

The ELM algorithm offers several advantages such as fast training speed, low computational complexity, and good prediction accuracy. The developed system also includes a user-friendly Streamlit interface that allows users to input agricultural parameters and receive instant yield predictions.
The proposed system can assist farmers, researchers, and policymakers in agricultural planning and decision-making**.** In the future, the system can be improved by incorporating larger datasets, real-time weather data, and advanced machine learning models to further enhance prediction accuracy.

ACKNOWLEDGMENT

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.13397

The authors would like to express their sincere gratitude to the Department of Computer Science and the project guide for their continuous guidance and support during the development of this research work.

REFERENCES

[1]. M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, “A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction,” IEEE Access, vol. 9, pp. 63406–63438, 2021.
[2]. M. Ashfaq, I. Khan, A. Alzahrani, M. U. Tariq, H. Khan, and A. Ghani, “Accurate wheat yield prediction using machine learning and climate-ndvi data fusion,” IEEE Access, vol. 12, pp. 40947–40961, 2024.
[3]. A. Mitra, S. Beegum, D. Fleisher, V. R. Reddy, W. Sun, C. Rayd, D. Timlind, and A. Malakar, “Cotton yield prediction: A machine learning approach with field and synthetic data,” IEEE Access, vol. 12, p. 101288, 2024.
[4]. Y. Alebele, W. Wang, W. Yu, X. Zhang, X. Yao, Y. Tian, Y. Zhu, W. Cao,and T. Cheng, “Estimation of crop yield from combined optical and sar imagery using gaussian kernel regression,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, p. 10520, 2021.
[5]. S. Yang, L. Hu, H. Wu, H. Ren, H. Qiao, P. Li, and W. Fan, “Integration of crop growth model and random forest for winter wheat yield estimation from uav hyperspectral imagery,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, p. 6253, 2021.
[6]. A. Badshah, B. Y. Alkazi, F. Din, K. Z. Zamli, and M. Haris, “Crop classification and yield prediction using robust machine learning models for agricultural sustainability,” IEEE Access, vol. 12, p. 162813, 2024.
[7]. K. Narayanasamy and I. Venkatachalam, “Predicting sugar yield from sugarcane using machine learning for jaggery production,” IEEE Access, vol. 13, p. 106090,2025.
[8]. S. Iniyan and R. Jebakumar, “Phenotype based smart mobile application for crop yield prediction and forecasting using machine learning and time series models,”p. 634, 2022.

