An Approach for Crop Prediction in Agriculture – Integrating Genetic Algorithm and Machine Learning

1Gattu Nikil, 2Bandari Bunny, 3Golla Prashanth, 4Konammagari Srikar, 5. Dr K Madan Mohan

1,2,3,4, Students, 5 Associate Professor

Department of Information Technology

GURU NANAK INSTITUTE OF TECHNOLOGY

ABSTRACT

Agriculture plays a vital role in the economy of many countries, especially in regions where a large population depends on farming. However, farmers face challenges such as unpredictable weather conditions, soil variability, and improper crop selection, leading to reduced productivity. This research proposes an intelligent crop prediction system using a combination of Machine Learning, Deep Learning, and Genetic Algorithm techniques. The system utilizes soil parameters such as Nitrogen, Phosphorus, Potassium, and pH along with environmental factors like temperature, humidity, and rainfall to predict suitable crops. A Long Short-Term Memory (LSTM) model is used to capture time-series patterns in weather data, while Random Forest is applied for classification. Genetic Algorithm is used to optimize model parameters and improve performance. The system also integrates Explainable AI techniques to enhance interpretability. Experimental results show improved accuracy and efficiency compared to traditional methods. This approach helps farmers make better decisions and increases agricultural productivity.

KEYWORDS

Crop Prediction, Machine Learning, LSTM, Genetic Algorithm, Agriculture, Deep Learning

1. INTRODUCTION

Agriculture is the backbone of many developing countries, where a large portion of the population depends on it for livelihood. However, farmers often face difficulties due to unpredictable climatic conditions, soil variations, and lack of proper decision-making tools. These challenges result in low crop yield and financial instability. With the advancement of Artificial Intelligence and Machine Learning, smart agriculture solutions are becoming more effective. Crop prediction systems can assist farmers by recommending suitable crops based on environmental conditions. This research focuses on developing an intelligent system that integrates Machine Learning, Deep Learning, and optimization techniques to improve crop prediction accuracy. The proposed system uses soil nutrients such as Nitrogen, Phosphorus, and Potassium along with weather parameters like temperature, humidity, and rainfall. A hybrid model combining LSTM, Random Forest, and Genetic Algorithm is used to provide accurate and reliable predictions.

2. LITREATURE REVIEW

1) Several research studies have explored crop prediction using machine learning techniques. Ramesh Medar et al. (2019) developed a crop recommendation system using **Decision Tree** and **Naïve Bayes** algorithms, achieving moderate accuracy in predicting suitable crops based on soil and climate conditions. Similarly, K. Shankar et al. (2020) applied **Support Vector Machines (SVM)** and **K-Nearest Neighbors (KNN)** for crop classification, but reported limitations when handling large and complex datasets.
2) With advancements in technology, researchers have shifted towards deep learning approaches. S. R. Dubey et al. (2021) proposed a **Multi-Layer Perceptron (MLP)** model for crop prediction, showing improved accuracy compared to traditional methods. In another study, A. Khaki and L. Wang (2019) utilized deep learning techniques for yield prediction, demonstrating the effectiveness of neural networks in agricultural forecasting.
3) The integration of IoT and real-time data has also been explored. S. Kamilaris and F. X. Prenafeta-Boldú (2018) highlighted the role of IoT and AI in smart agriculture, emphasizing improved prediction accuracy through real-time environmental monitoring.
4) Recent studies also focus on explainable AI for better transparency. Marco Tulio Ribeiro et al. (2016) introduced **LIME**, and Scott Lundberg and Su-In Lee (2017) proposed **SHAP**, both widely used for interpreting machine learning models in agriculture and other domains.
5) However, many existing systems fail to capture temporal dependencies in agricultural data. To address this, Sepp Hochreiter and Jürgen Schmidhuber (1997) introduced **Long Short-Term Memory (LSTM)** networks, which are effective for time-series prediction. Additionally, John Holland (1975) proposed the **Genetic Algorithm (GA)** for optimization problems. This research builds upon these approaches by integrating LSTM with Genetic Algorithm to improve prediction accuracy and model performance in crop prediction systems.

3. PROPOSED METHODOLOGY

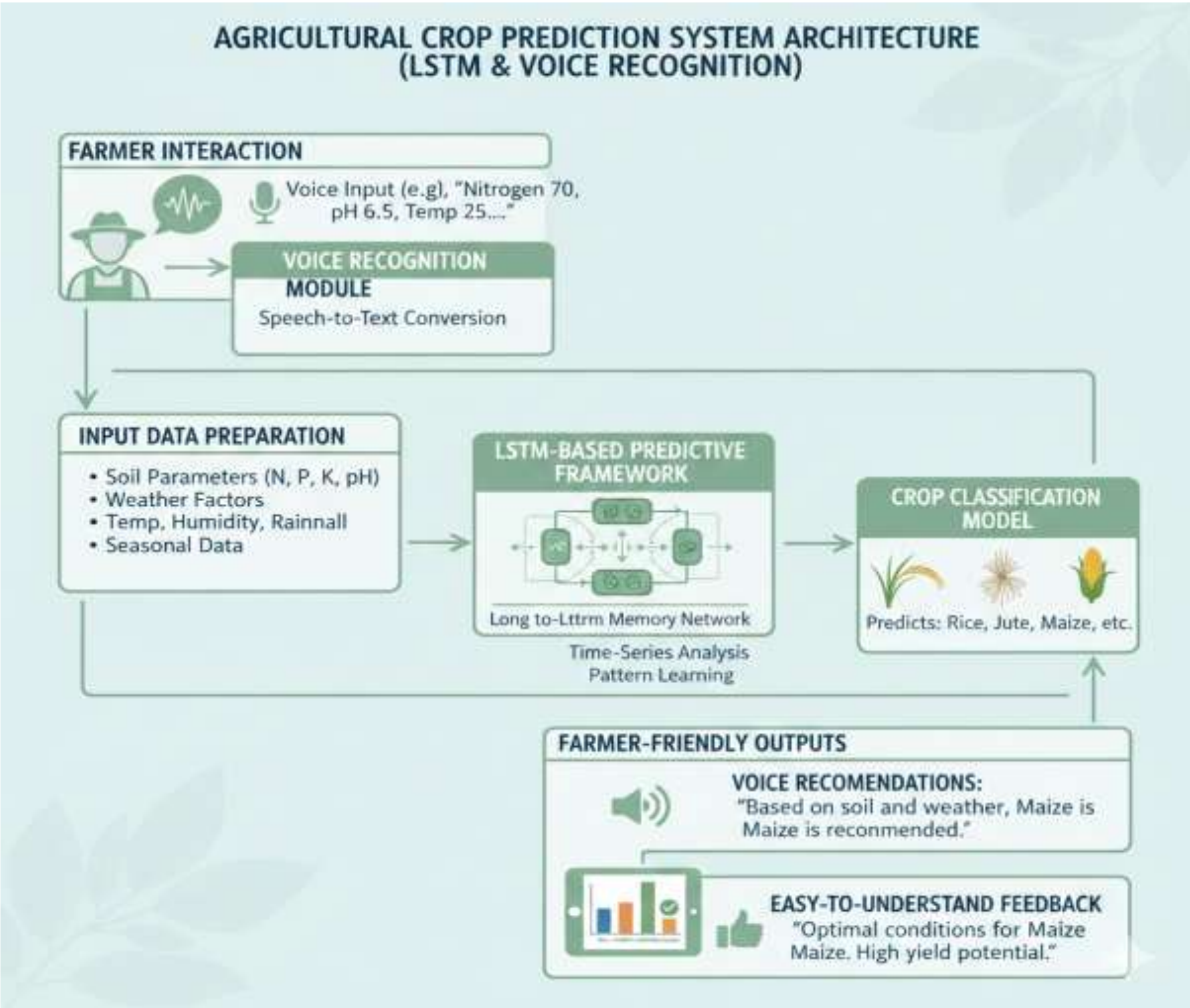
The proposed system follows a structured approach for crop prediction using multiple techniques. Initially, a dataset containing soil and weather parameters is collected. The dataset includes features such as Nitrogen, Phosphorus, Potassium, pH, temperature, humidity, and rainfall.

The data is preprocessed by handling missing values, normalization, and feature scaling. After preprocessing, the dataset is divided into training and testing sets.

A hybrid model is implemented using LSTM, Random Forest, and Genetic Algorithm. The LSTM model is used to analyze time-series weather data and capture temporal patterns. Random Forest is applied for classification of crop types based on input features. Genetic Algorithm is used to optimize hyperparameters such as learning rate, number of layers, and estimators to improve accuracy.

The system also integrates Explainable AI techniques like LIME and SHAP to interpret model predictions. This helps in understanding the influence of different features on crop selection. Finally, the system predicts the most suitable crop and provides additional recommendations such as fertilizer usage and yield estimation.

3.1 SYSTEM ARCHITECTURE



Step 1: Farmer Interaction

- The process starts with the farmer.
- The farmer provides input using voice instead of typing.

- Example input:
  - Nitrogen 70, pH 6.5, Temperature 25
- Makes the system easy to use, especially for rural farmers.

Step 2: Voice Recognition Module

- The voice input is processed using a speech-to-text system.
- It converts spoken data into text format.

  Example:
- Voice → Nitrogen 70
- Converted to → N = 70
- Enables hands-free and user-friendly interaction

Step 3: Input Data Preparation

- The converted data is combined with:
  - Soil parameters → Nitrogen (N), Phosphorus (P), Potassium (K), pH
  - Weather data → Temperature, Humidity, Rainfall
  - Seasonal data
- Data is then:
  - Cleaned
  - Normalized
  - Prepared for model input
- Ensure accurate and consistent data for prediction

Step 4: LSTM-Based Predictive Framework

- The prepared data is given to the LSTM (Long Short-Term Memory) model.

What LSTM does:

- Analyzes time-series data (weather patterns over time)
- Learns:
  - Seasonal changes
  - Climate patterns
  - Long-term dependencies
- Improves prediction accuracy by understanding past trends

Step 5: Crop Classification Model

- After LSTM processing, the system uses a classification model.

It predicts:

- Suitable crops like:

- Rice
- Maize
- Jute

- Final decision: Which crop is best for given conditions

Step 6: Farmer-Friendly Outputs

The system gives results in an easy way:

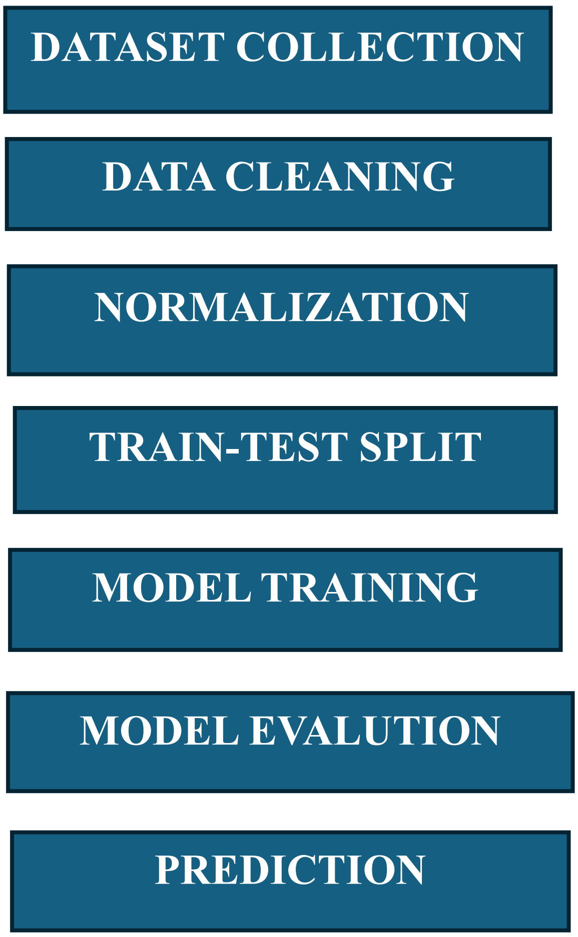
Voice Recommendations

- Output is given as audio
- Example:
  - Maize is recommended based on your soil and weather

Visual Feedback

- Charts and simple messages are shown
- Example:
  - High yield potential for Maize
- Easy understanding for farmers

3.2 MODEL WORKFLOW DIAGRAM

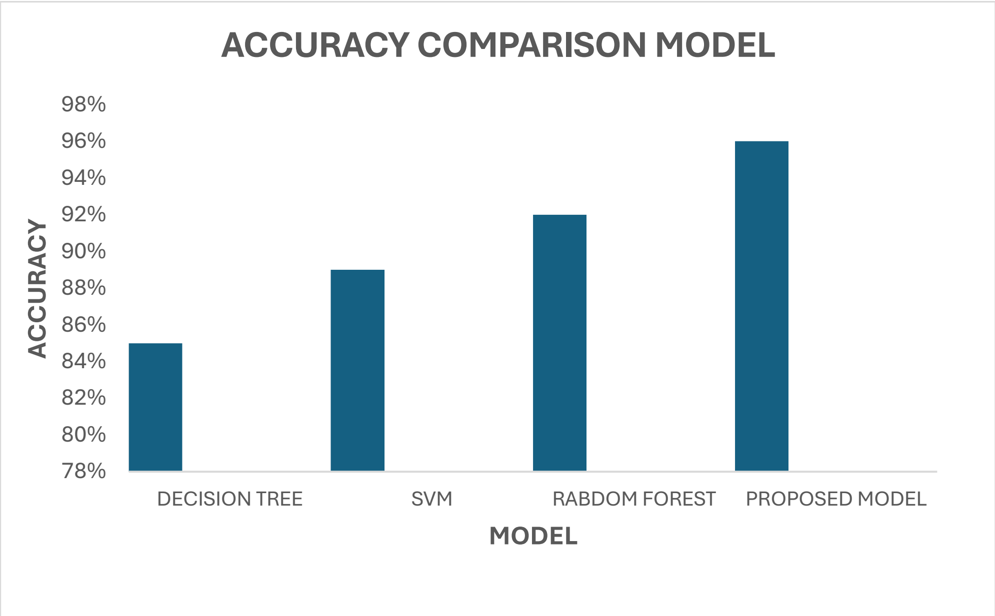


RESULTS AND DISCUSSION

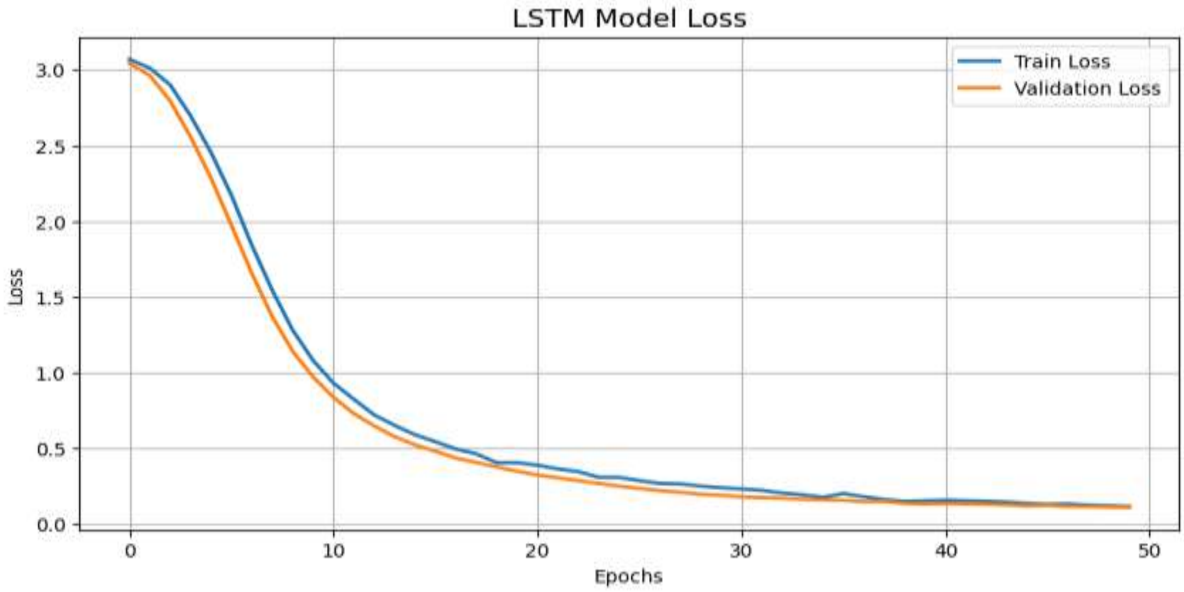
- The proposed model was tested using a dataset containing multiple crop types and environmental parameters. The performance of the model was evaluated using metrics such as accuracy, precision, recall, and F1-score.
- The hybrid model combining LSTM, Random Forest, and Genetic Algorithm achieved higher accuracy compared to traditional machine learning models. The use of LSTM improved the ability to capture seasonal variations, while Genetic Algorithm enhanced model optimization.
- The system also provides user-friendly outputs, including recommended crops, expected yield, fertilizer suggestions, and pest control measures. The integration of Explainable AI techniques helped in understanding the model's decision-making process. Overall, the results demonstrate that the proposed system is efficient, accurate, and practical for real-world agricultural applications.

3.2 ACCURACY COMPARISON GRAPH

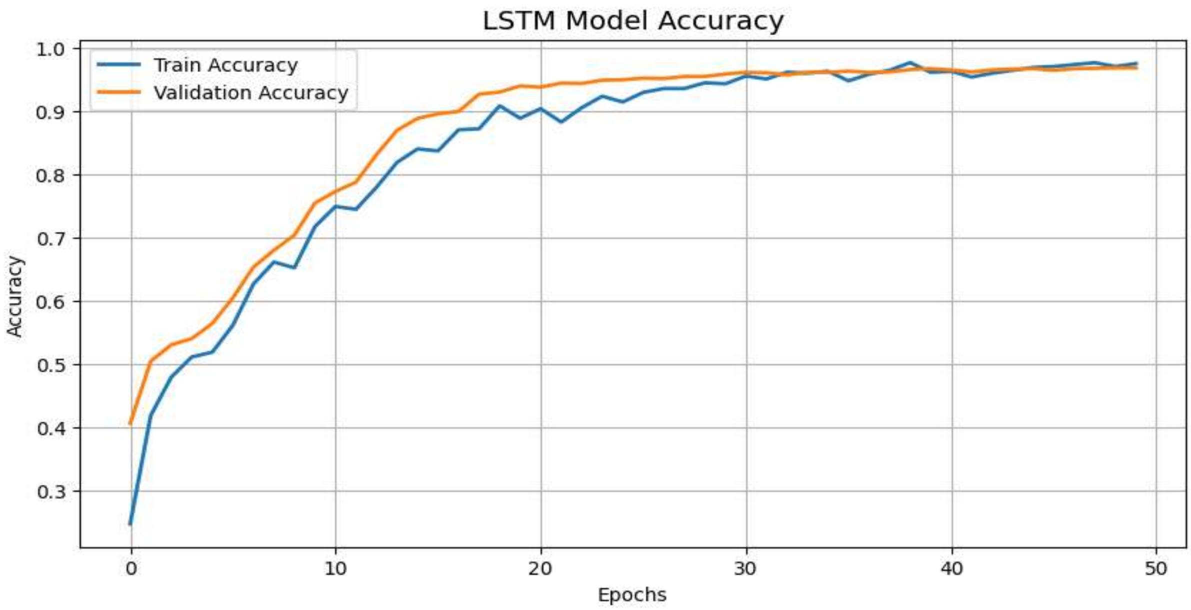
MODEL	ACCURACY
Decision Tree	85%
SVM	89%
RANDOM FOREST	92%
PROPOSED MODEL	96%



4.2 LSTM MODEL LOSS

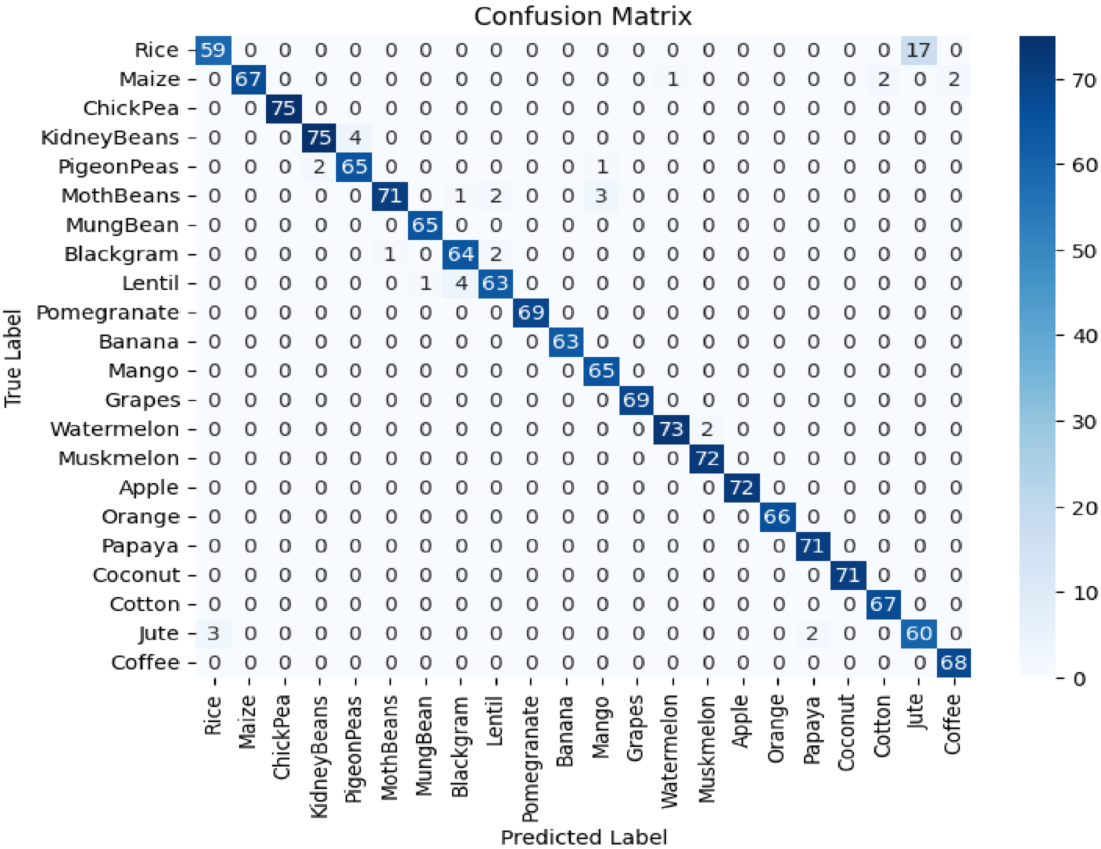


4.2 LSTM MODEL ACCURACY



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 04 | April - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

4.3 CONFUSION MATRIX



4. CONCLUSION

- This research presents an intelligent crop prediction system using Machine Learning, Deep Learning, and Genetic Algorithm techniques. The system effectively analyzes soil and weather conditions to recommend suitable crops. The use of LSTM helps in capturing time-dependent patterns, while Genetic Algorithm improves model performance through optimization.
- The proposed approach provides accurate predictions and supports farmers in making informed decisions. It also enhances agricultural productivity and reduces risks associated with crop selection.The system also integrates Explainable AI techniques to improve transparency and reliability. Additionally, the voice recognition feature makes the system user-friendly and accessible to farmers.
- Experimental results show that the proposed model performs better than traditional methods. The system helps farmers make informed decisions and increases productivity. Overall, this approach supports smart and sustainable agriculture. Future improvements can include real-time data integration and large-scale deployment.

5. REFERENCES

[1] M. H. Imam, N. Nahar, R. Bhowmik, S. B. S. Omit, T. Mahmud, M. S. Hossain, and K. Andersson, "A transfer learning-based framework: Mobilenet-svm for efficient tomato leaf disease classification," in 2024 6th International Conference on Electrical Engineering and Information &Communication Technology (ICEEICT), pp. 693–698, 2024.

[2] T. Mahmud, K. Barua, A. Barua, N. Basnin, S. Das, M. S. Hossain, and K. Andersson, "Explainable ai for tomato leaf disease detection: Insights into model interpretability," in 2023 26th International Conference on Computer and Information Technology (ICCIT), pp. 1–6, IEEE, 2023.

[3] T. Mahmud, M. Ptaszynski, and F. Masui, "Exhaustive study into machine learning and deep learning methods for multilingual cyberbullying detection in bangla and chittagonian texts," Electronics, vol. 13, no. 9, p. 1677, 2024.

[4] S.UmmeHabiba, F.Tasnim, M.S.Hasan Chowdhury, M.K.Islam, L.Nahar, T. Mahmud, M. S. Kaiser, M. S. Hossain, and K. Andersson, "Early prediction of chronic kidney disease using machine learning algorithms with feature selection techniques," in International Conference on Applied Intelligence and Informatics, pp. 224–242, Springer, 2023.

[5] L. Bi and G. Hu, "A genetic algorithm-assisted deep learning approach for crop yield prediction," Soft Computing, vol. 25, pp. 10617–10628, 2021.

[6] V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, and A. Janani, "An effective crop prediction using random forest algorithm," in 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1–5, IEEE, 2020.

[7] N. Jain, A. Kumar, S. Garud, V. Pradhan, and P. Kulkarni, "Crop selection method based on various environmental factors using machine learning," International Research Journal of Engineering and Technology (IRJET),vol. 4, no. 02, pp. 551–558, 2017.

[8] S. Jain and D. Ramesh, "Machine learning convergence for weather based crop selection," in 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–6, IEEE, 2020.