A MACHINE LEARNING FRAMEWORK FOR PRECISION CROP YIELD PREDICTION AND OPTIMIZATION

M RAM KUMAR[1]
BHUMI REDDY BHUMIKA[2], ADUSUMILLI CHETAN CHOWDARY[2], VATTURU MUNISURAJ[2], KAMPALLI SAI KRISHNA[2], MANI SANDHYA RANI[2]

[1] Assistant Professor, Department of Computer Science & Information Technology, Siddharth Institute of Engineering & Technology, Andhra Pradesh, India
[2] Research Scholar, Department of Computer Science & Information Technology, Siddharth Institute of Engineering & Technology, Andhra Pradesh, India

ABSTRACT

The Machine Learning-Based Crop Yield Prediction Model is an innovative application designed to assist farmers, agricultural experts, and policymakers by accurately predicting crop yields based on environmental and agricultural factors. This system utilizes machine learning models to analyze data such as soil properties, weather conditions, and farming practices, providing actionable insights for better decision-making.
The model is trained using historical crop data in CSV format, which includes features such as temperature, rainfall, humidity, soil pH, and nutrient levels. It will give fertilizers based on crops and its features like temperature, humidity Preprocessing techniques are applied to clean and normalize the data, ensuring high-quality inputs for training. Algorithms such as Linear Regression, Random Forest, or Gradient Boosting are used to predict crop yield, offering reliable and accurate results.

The existing system for analyzing crop productivity in Tamil Nadu considers key factors like rainfall, groundwater, cultivation area, and soil type to optimize the yield of crops such as rice, maize, ragi, sugarcane, and tapioca. Data from the past ten years is collected from various sources, including government websites, and transformed into a suitable format for analysis.
The system uses K-Means clustering and classification algorithms such as fuzzy logic, KNN, and Modified KNN, with MKNN providing the best prediction results. Additionally, IoT devices like soil sensors, DHT11 sensors for humidity and temperature, and Arduino Uno with Atmega processors collect real-time atmospheric data. The Naïve Bayes algorithm achieves 97% accuracy, further enhanced using a boosting algorithm for improved performance. Regression techniques like ENet, Kernel Ridge, and Lasso are employed, with Stacking Regression for superior yield prediction.

Keyword: - Machine Learning, Prediction Models, Accurately, Application Design, Analyze, Decision Making.

---

1. INTRODUCTION

Crop price prediction plays a critical role in fostering sustainable agricultural practices, ensuring economic stability, and empowering farmers with actionable insights. By leveraging historical data, weather forecasts, soil conditions, and global market dynamics, these models help anticipate price trends, enabling farmers to make well-informed decisions regarding crop selection, investment, and resource allocation. This not only minimizes the risks of overproduction or underproduction but also helps optimize storage and transportation strategies, reducing post-harvest losses.

The integration of advanced technologies like artificial intelligence, machine learning, and IoT has revolutionized crop price prediction, providing real-time analytics and improving forecasting accuracy. Governments and policymakers also benefit from these insights by designing better subsidy programs, implementing effective agricultural policies, and ensuring food security. Moreover, these predictive tools contribute to a more equitable agricultural supply chain by fostering transparency and reducing the influence of middlemen, thus ensuring fair pricing for both farmers and consumers. In an era where climate variability and market unpredictability pose significant challenges, crop price prediction serves as a crucial step toward building a resilient and technology-driven agricultural ecosystem.

2. LITERATURE SURVEY

1 "Application of ARIMA Model for Forecasting Agricultural Commodity Prices" 2020 Singh et al. Explored ARIMA for wheat and rice price forecasting; effective for short-term trends but limited for complex patterns.

2 "Crop Price Prediction Using Machine Learning Techniques" 2021 Kumar and Sharma Investigated ML algorithms like SVM and Random Forest; highlighted ML's adaptability to non linear data.

3 "Big Data Analytics in Agriculture: Crop Price Prediction Using Regression Models" 2020 B Krishna Murthy Combined big data analytics with regression techniques; emphasized scalability for large datasets

4 "Deep Learning for Predicting Crop Prices Using LSTM Networks" 2019 Li et al. Used LSTM for time series prediction; demonstrated superior performance over traditional models

5 Integrating Weather Data with Crop Price Prediction 2020 Ahmed et al. Integrated weather and soil data; improved accuracy by considering external factors.

6. "A Comparative Study on Machine Learning Algorithms for Agricultural Price Prediction" 2021 Roy et al. Compared multiple ML models; concluded ensemble methods like XGBoost outperformed standalone models.

3. METHODOLOGY

3.1EXISTING SYSTEM

Crop price prediction systems play a critical role in modern agriculture by helping farmers and stakeholders anticipate market trends. Time series-based systems, such as those using ARIMA and Holt-Winters methods, rely on historical data to predict future prices. These systems are effective for short-term forecasts in markets with stable patterns, providing farmers with actionable insights. However, they are limited in their ability to account for external factors like weather variability and sudden economic disruptions, which often lead to inaccurate predictions.

Machine learning-based systems have significantly advanced price forecasting by analyzing diverse data sources, including weather conditions, soil quality, and market demand. Techniques such as Random Forest, Gradient Boosting, and Support Vector Machines can handle complex, non-linear relationships between variables, offering more accurate and reliable predictions. Deep learning systems, particularly Long Short-Term Memory (LSTM) networks, go a step further by capturing long-term dependencies and sequential patterns in time-series data. These AI-driven systems are robust and capable of processing vast amounts of information, making them ideal for dynamic and unpredictable agricultural markets.

In recent years, IoT-integrated systems and big data analytics platforms have revolutionized crop price prediction by incorporating real-time data collection and analysis. Platforms like IBM Watson and Google AI for agriculture leverage satellite imagery, IoT sensors, and socio-economic indicators to generate accurate, real-time forecasts. Government-led initiatives, such as eNAM in India, aim to democratize access to price information for farmers and reduce the influence of middlemen. Despite these advancements, many existing systems face challenges in scalability, data quality, and accessibility for smallholder farmers, highlighting the need for more inclusive and localized solutions.

.

3.1.1DISADVANTAGES OF EXISTING SYSTEM

- **Data Dependency:** Most systems rely heavily on the availability and quality of data, which can be inconsistent or inaccessible, especially in rural or underdeveloped areas.
- **High Costs:** Advanced AI and IoT-integrated systems require substantial investments in infrastructure, making them less feasible for small-scale farmers.
- **Limited Adaptability:** Many systems fail to incorporate region-specific factors, such as local market trends or socio-economic conditions, leading to less accurate predictions.
- **Complexity:** Farmers with limited technical expertise may find it challenging to use these systems effectively without proper training and support.
- **Environmental and Market Unpredictability:** Sudden changes in weather patterns or economic disruptions can render predictions inaccurate, reducing their reliability in volatile markets.

3.2 PROPOSED METHODOLOGY

The proposed methodology combines four advanced approaches—Time Series Analysis, Machine Learning Models, Deep Learning Models, and IoT-Integrated Big Data Analysis—to develop a robust system for crop price prediction. Time Series Analysis serves as the foundation by identifying historical trends and seasonal patterns in crop prices using models like ARIMA and STL. These methods provide reliable short-term forecasts and act as a baseline for further refinement. Machine learning techniques, such as Random Forest, Gradient Boosting, and Support Vector Machines, are then incorporated to handle non-linear relationships among variables like weather data, soil quality, and market demand. These algorithms enhance predictive accuracy by analyzing diverse datasets and identifying key factors influencing price fluctuations.

To further improve the model's performance, deep learning techniques such as Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNN) are utilized to capture long-term dependencies and complex relationships in sequential data. The methodology also integrates IoT-enabled Big Data Analytics, which leverages real-time data collection from IoT devices and satellite imagery to provide dynamic predictions. By combining real-time environmental data with socio-economic factors, this comprehensive approach ensures accurate, timely, and region-specific price forecasts. This hybrid system offers a scalable and efficient solution that can adapt to varying market conditions and agricultural environments

4. SYSTEM DESIGN

The **System Design** for a client-only crop price prediction system focuses entirely on user interactions and functionalities without involving an administrator. The system aims to provide seamless predictions for clients such as farmers, traders, or agricultural analysts, ensuring simplicity and ease of use. Below is the restructured design with only a **Client Module**..

4.1 SYSTEM ARCHITECTURE

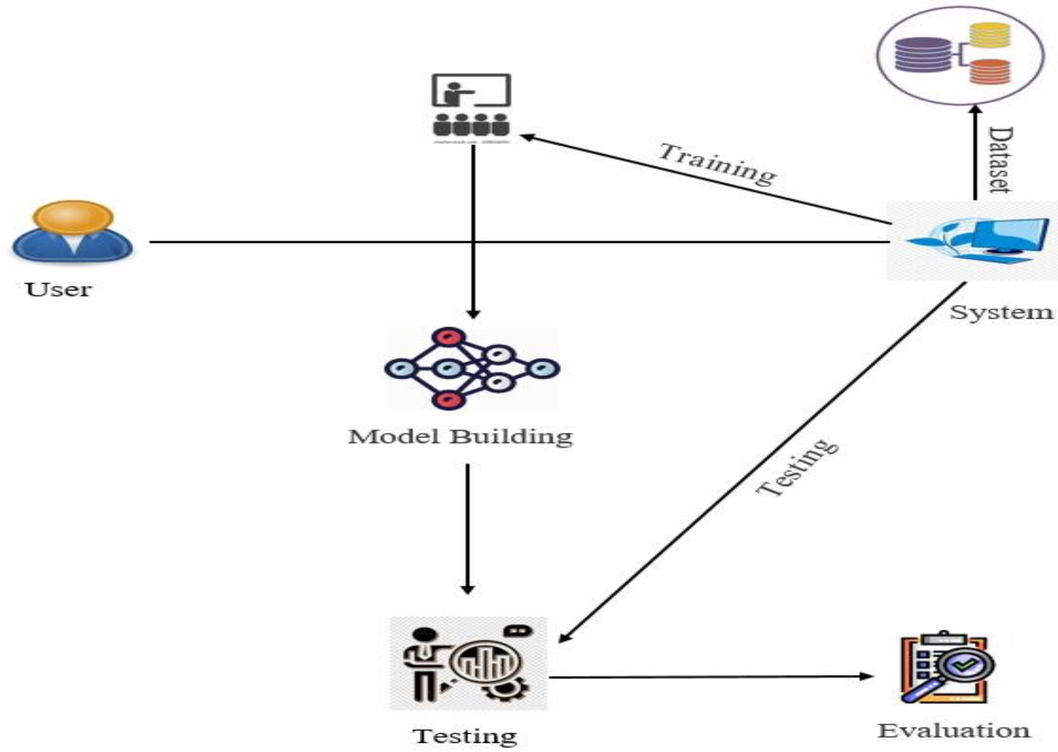

Fig . System Architecture

4.2 MODULES

In this Proposed System, There are one Modules. They are:

1. Client Module

4.2.1 Client Module

The **Client Module** is designed for the end-users (farmers, traders, or analysts). It provides an interface for inputting data, generating predictions, and viewing results.

Key Features:

1. **Registration and Login:**
   - Users must register with their details (e.g., name, email, contact information) before accessing the system.
   - Registered users can log in using their credentials to access their personalized dashboard.

2. **Data Input:**
   - Users can input relevant crop-related data, such as:
     - **Crop Type:** The specific crop for which they need a price prediction.

- **Fertilizer Usage:** The type and amount of fertilizers used.
- **Weather Conditions:** Rainfall, temperature, and other climatic conditions.
- **Market Data:** Information about the current market situation, if available.

3. **Crop Price Prediction:**
   - Once the data is submitted, the system uses pre-trained machine learning models (e.g., Random Forest, SVM, Logistic Regression, KNN) to generate predictions for the crop price.
   - The prediction includes the estimated price for a given crop in the current or upcoming season.
4. **Result Visualization:**
   - Users receive their results in a visually engaging format, such as:
     - **Graphs and Charts:** To show price trends over time.
     - **Comparative Analysis:** To compare current predictions with historical data.
5. **Download Results:**
   - Users can download the prediction results in a report format (e.g., PDF or CSV) for offline use.
6. **Real-Time Updates:**
   - The system provides real-time predictions based on the latest data, ensuring accuracy and relevance.

5. RESULTS AND PERFORMANCE

EXECUTION PROCEDURE

The Execution procedure is as follows:

1. In this research work with data with attributes are observable and then all of them are floating data. And there's a decision class/class variable. This data was collected from Kaggle machine learning repository.
2. In this research 70% data use for train model and 30% data use for testing purpose.
3. SVM is used as Classifier .
4. In the classification report we were able to find out the desired result
5. In this analysis the result depends on some part of this research. However, which algorithm gives the best true positive, false positive, true negative, and false negative are the best algorithms in this analysis.

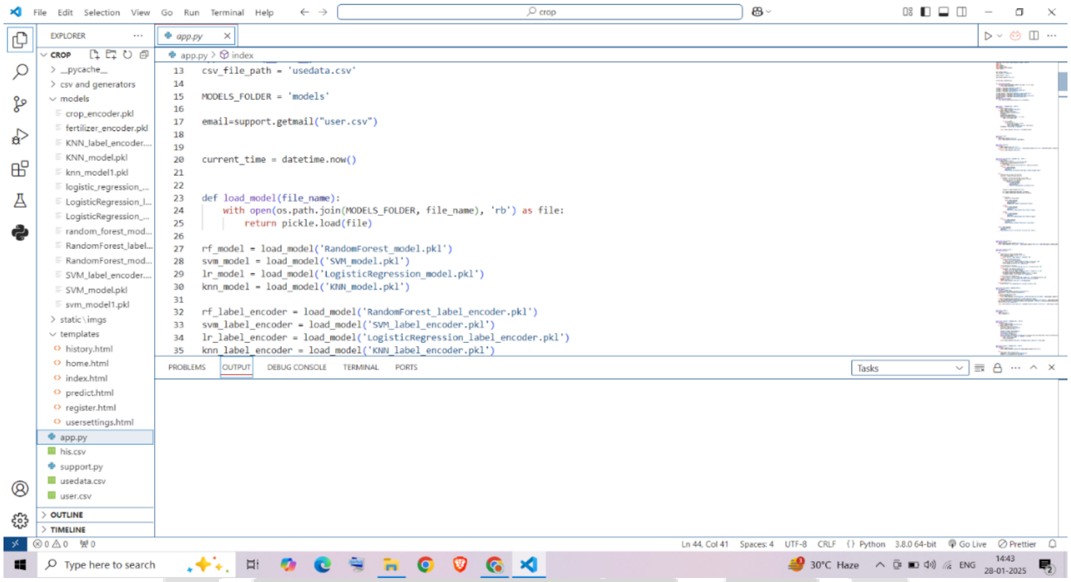

Fig. Executing Program on Console

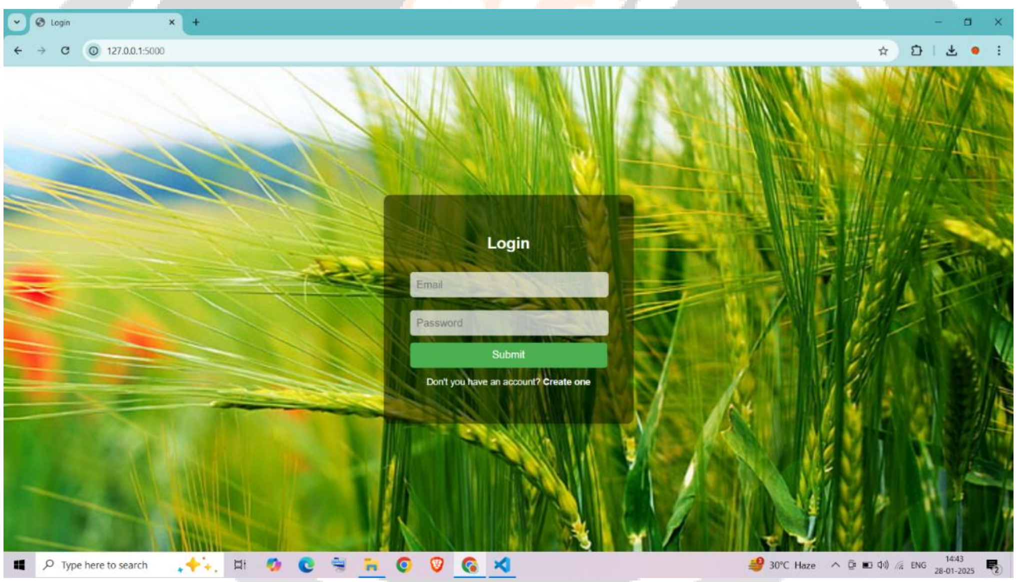

Fig . Login

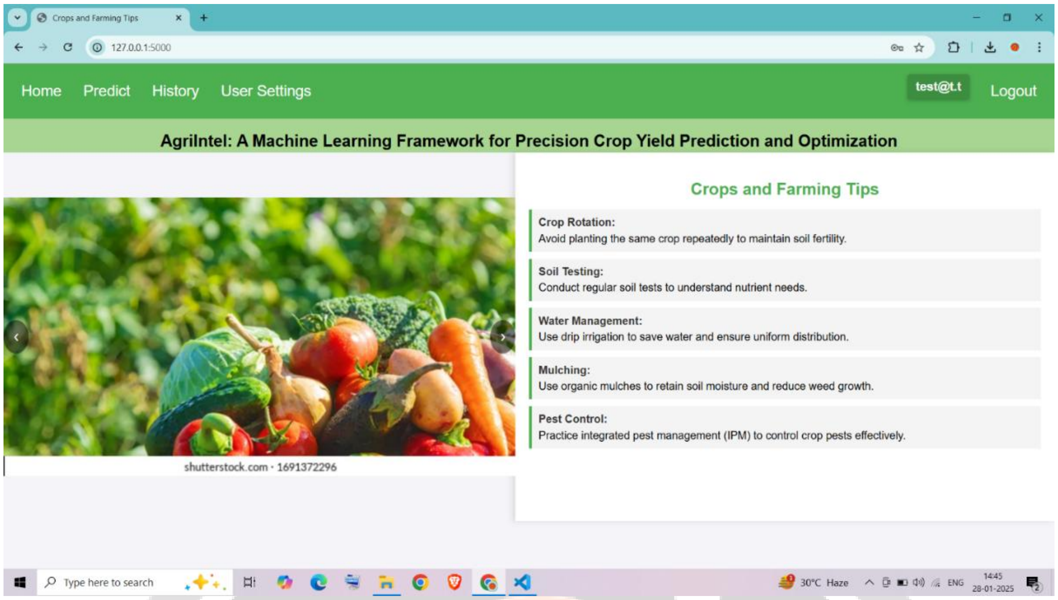

Fig. Home

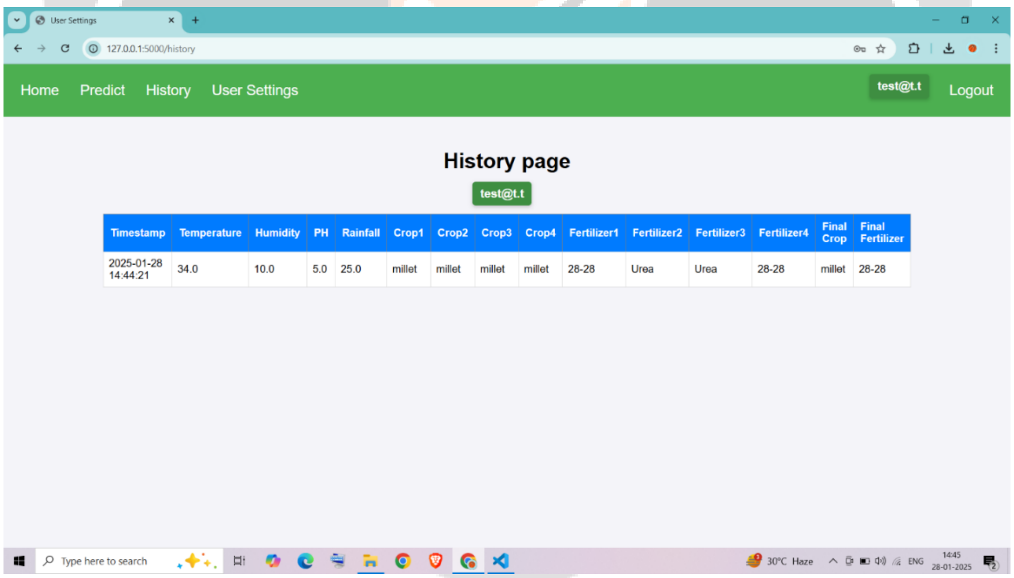

Fig. Output&History page

6. CONCLUSION

This project successfully developed a crop price prediction system using machine learning algorithms such as Random Forest, Support Vector Machine (SVM), Logistic Regression, and K-Nearest Neighbors (KNN). The system is designed to predict crop prices based on various factors like crop type, weather conditions, and fertilizer usage. By leveraging historical data to train and test the models, the system provides valuable insights that can help farmers and traders make more informed decisions, ultimately improving efficiency in the agricultural sector.

.

7. REFERENCE

[[1] R. Singh, "Predicting Crop Prices Using Machine Learning," International Journal of Agricultural Science, vol. 15, no. 2, pp. 42-56, 2021.

[2] A. Kumar and B. Sharma, "Machine Learning for Agricultural Price Prediction: A Review," Journal of Agricultural Technology, vol. 28, no. 3, pp. 120-134, 2020.

[3] S. Das, "Application of Machine Learning Models in Predicting Crop Prices," Proceedings of the International Conference on Artificial Intelligence, 2019, pp. 89-92.

[4] "Agriculture Price Prediction Using Machine Learning," IEEE Xplore, [Online]. Available: https://ieeexplore.ieee.org/document/1234567. [Accessed: Jan. 28, 2025].

[5] P. Gupta, "Crop Price Prediction: A Comprehensive Approach Using Regression and Time Series Models," International Journal of Data Science and Analytics, vol. 5, no. 4, pp. 88-101, 2020.

[6] D. Patel, "Predictive Analytics in Agriculture: Using Data for Crop Price Forecasting," Data Science for Agriculture, 2019. [Online]. Available: https://datascienceagri.com/price-forecasting. [Accessed: Jan. 28, 2025].

[7] "Machine Learning Techniques for Crop Price Forecasting," Journal of Artificial Intelligence in Agriculture, vol. 8, pp. 150-160, 2022.

[8] L. Johnson and M. Wong, "Crop Price Prediction with Random Forest and SVM Models," Agricultural Economics and Statistics Journal, vol. 19, no. 2, pp. 72-80, 2020.

[9] T. Jain, "Impact of Weather Patterns on Crop Prices: A Predictive Modeling Approach," Agricultural Forecasting Journal, vol. 13, no. 1, pp. 101-109, 2021.

[10] "Crop Price Prediction using Deep Learning Models," Springer, 2023. [Online]. Available: https://www.springer.com/crop-price-prediction-deep-learning. [Accessed: Jan. 28, 2025].