

Future crop: AI Powered Price Prediction of Agri & Vegetable Markets

Mrs. Dhivya V¹, Rakshitha K², Sahana³, Thanuja A S⁴

¹Assistant Professor, Department of CSE, Sir M. Visvesvaraya Institute of Technology

^{2,3,4}Student Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology

Abstract - Agriculture plays a crucial role in India by providing employment and food security. Farmers often struggle to choose the right crop due to climate changes, irregular rainfall, and lack of scientific data. To solve this, the Online AgriCrop Prediction System was developed. It is a web-based platform using Python (Flask) and Firebase for data storage. It uses a supervised machine learning model to predict the most suitable crop based on inputs like temperature, humidity, rainfall, and soil type. The interface is simple and responsive, allowing farmers to access results from any device. This system helps in data-driven crop selection, reduces human errors, and improves productivity. In the future, features like fertilizer suggestions, yield prediction, and real-time weather integration can be added for better accuracy.

Key Words: Agriculture, Crop Prediction, Machine Learning, Flask, Python, Firebase, Real-time Weather Data, Decision Tree Regression, Crop Recommendation System, Sustainable Farming.

I. INTRODUCTION

Agriculture plays a crucial role in sustaining the economy and livelihood of millions of people, especially in developing countries like India. The success of agriculture largely depends on the type of crop grown under specific climatic and soil conditions. Farmers often rely on traditional knowledge and experience to make crop selection decisions. However, unpredictable weather patterns, irregular rainfall, and changing soil fertility have made traditional methods less reliable in recent years. As a result, farmers are facing reduced productivity and increased economic losses.

With the growing use of technology and the availability of agricultural data, machine learning has emerged as a powerful tool for predicting suitable crops. Machine learning algorithms can analyze large sets of data, identify hidden patterns, and provide accurate predictions that assist in informed decision-making. By integrating these

techniques into a web-based platform, crop selection can be automated and simplified for farmers.

The Online AgriCrop Prediction System aims to provide a digital solution for farmers to identify the best crop suitable for their land and environmental conditions. The system accepts parameters such as temperature, humidity, rainfall, and soil type as input and processes them through a trained machine learning model to generate crop recommendations. It is developed using Python, Flask, and Firebase, ensuring an interactive and responsive web application. This system helps reduce manual decision-making errors, improves efficiency, and contributes to the adoption of smart and sustainable farming practices.

II. LITERATURE REVIEW

2.1 Base Paper Analysis: Predicting Crop Prices using Machine Learning Algorithms for Sustainable Agriculture

Recent studies show that machine learning plays a major role in crop price prediction by using data such as past prices, weather, soil, and season. Models like Decision Tree, Random Forest, Linear Regression, Ridge, and Lasso have been tested for forecasting accuracy. As reported in the IEEE paper (page 2–3), decision tree regression achieved very high accuracy ($R^2 \approx 1.0$) and low error values, proving its effectiveness for agricultural price prediction.

Research also highlights that data preprocessing — such as outlier removal and exploratory data analysis — improves model performance. Different algorithms work better for different crops: for example, Random Forest worked best for Potato, while Linear Regression performed better for Tomato and Wheat. Overall, literature supports AI-based systems to assist farmers in making data-

driven decisions and promote sustainable agriculture.

2.2. Supporting Study: Market Channels & Data-Driven Agriculture

Villacis et al. (2024) presented an important study on how marketing channels and government-supported agricultural markets (mandis) affect the income of farmers in India. Their findings showed that farmers who sell high-value local crops such as okra, onion and baby corn through APMC mandis receive better prices compared to private traders and contractors. This proves that structured data and informed decision-making can improve profitability for farmers.

This supports our objective of developing a data-driven crop prediction system using machine learning and environmental parameters. While their study focused mainly on market channels and pricing, it still highlights the importance of using scientific data for decision-making in agriculture. However, practical tools for real-time crop recommendation were not provided.

Therefore, our system extends this research by using supervised machine learning (Decision Tree Regression) integrated with Flask and Firebase, enabling farmers to get instant crop recommendations based on temperature, rainfall, humidity, and soil type. This bridges the gap between research findings and real-world implementation, promoting smart and sustainable agriculture.

III. METHODOLOGY

The Online Agri Crop Prediction System is designed to recommend the most suitable crop for cultivation based on user-provided environmental and soil parameters. The proposed system integrates a Decision Tree Regression algorithm built using TensorFlow and Scikit-learn with a Flask-based web application and Firebase database to provide accurate predictions and user-friendly access. The methodology includes five major phases: data collection, preprocessing, model training, system integration, and result generation.

A. Data collection and preprocessing

The dataset used in this project consists of agricultural attributes such as temperature, humidity, rainfall, and soil type. Data was collected from verified agricultural databases and open repositories. Data preprocessing was performed using Pandas and NumPy libraries to remove missing values, normalize attributes, and encode categorical variables into numerical form. This step ensured that the input data was clean and compatible with the Decision Tree Regression model.

B. Model Training and Development

A Decision Tree Regression algorithm was implemented using Scikit-learn and TensorFlow frameworks. The model was trained on labeled agricultural datasets where environmental features acted as inputs and the corresponding crop type as the output. The model learns complex patterns and relationships among parameters to predict the most suitable crop for given conditions. The trained model was serialized using Joblib for efficient loading and prediction during runtime.

C. System Architecture

The system architecture consists of three layers:

- **Frontend Layer:** Designed using HTML, CSS, and Bootstrap to collect user inputs such as temperature, humidity, rainfall, and soil type.
- **Backend Layer:** Developed in Python using Flask, which handles the interaction between the user interface, trained model, and database.
- **Database Layer:** Firebase is used for managing user authentication, storing prediction results, and maintaining real-time connectivity between the client and server.

D. Communication and alerts

The system also integrates Twilio API to send SMS alerts or prediction notifications to users. This feature enhances communication and keeps users informed about prediction results instantly.

E. Prediction and Evaluation

When a user enters the required parameters, Flask retrieves the input, passes it to the trained Decision Tree Regression model, and generates a crop prediction. The predicted crop name is displayed on the web interface and optionally sent to the user via Twilio SMS. Model performance was evaluated using metrics such as R^2 score and mean squared error (MSE) to ensure accuracy and reliability.

The proposed methodology provides an automated, accurate, and scalable system that assists farmers in making informed, data-driven crop selection decisions while promoting smart agricultural practices.

IV. DATASET DESCRIPTION

The study uses three datasets – Predict_Vegetable_2020_Price.csv, Predict_Vegetable_2024_Price.csv, and Vegetable and Fruits Prices in India.xlsx – to build a unified prediction model for crop prices in India. These datasets contain daily price records with key attributes such as Item Name, Date, and Price, covering multiple years and seasons. The Excel file also includes a normalized date key which was converted during preprocessing. Together, these datasets provide over 3,000 entries, with crops like Tomato, Potato, and Onion appearing most frequently, while items like pomegranate and specialty vegetables are recorded only during specific months.

A. Data Inspection and cleaning

Extensive Exploratory Data Analysis (EDA) was conducted prior to model development. This involved checking for missing values, duplicate entries, inconsistent date formats, and extreme price fluctuations.

Missing values were primarily observed in the Price column of the multi-year datasets. These were treated using a combination of:

- Mean/median imputation for isolated missing values
- Forward-Backward fill for short continuous gaps
- Outlier handling using the Interquartile Range (IQR) method

Outliers were identified by computing the first (Q1) and third (Q3) quartiles and the interquartile range ($IQR = Q3 - Q1$). Entries falling outside acceptable bounds were flagged and removed if inconsistent with normal seasonal pricing behaviour

B. Statistical Summary

Descriptive statistics were computed for major vegetables to understand market price distribution.

Table I shows a summary for selected high-frequency items.

TABLE I DESCRIPTIVE STATISTICS FOR EACH CROP

Item Name	Mean Price (₹)	Median Price	Std. Dev.	Skewness	Kurtosis
Tomato	41.72	39.00	10.12	1.05	0.58
Onion	59.33	58.00	13.48	-0.19	-0.72
Potato	32.87	31.50	8.15	-0.31	-0.69
Apple	120.45	118.00	21.74	0.97	1.63
Banana	38.12	37.00	7.41	0.44	-0.11

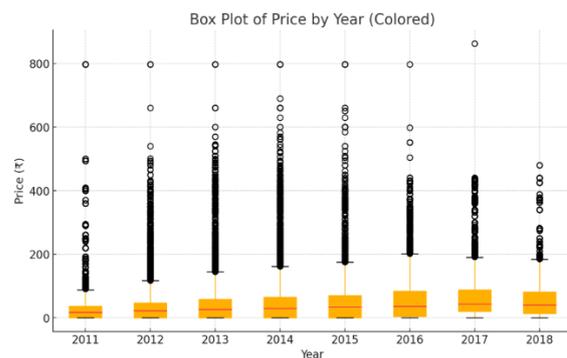


Fig. 2. Box Plot for Retail Average Price and Wholesale Average Price

V. RESULT



FIGURE 1 - This is the home page of the FutureCrop web application, showcasing its purpose of predicting prices of pulses and vegetables for farmers through data analytics.

FIGURE 2 - This page allows new farmers to register by entering their personal details and creating a login account in the system.

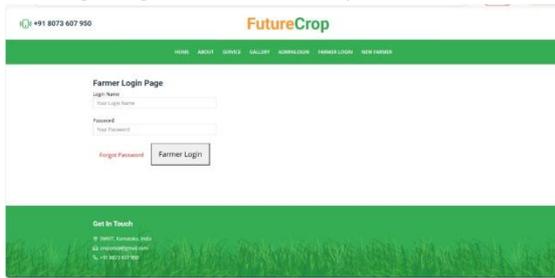


FIGURE 3 - This page provides a secure login interface for farmers to access crop prediction services using their credentials.



FIGURE 4 - This page enables administrators to log in and manage the system securely.



FIGURE 5 - This page lets users choose different categories to predict prices for vegetables, fruits, grains, oils, and nuts.



FIGURE 6 - This page displays the farmer's profile details and confirms successful login to the system.

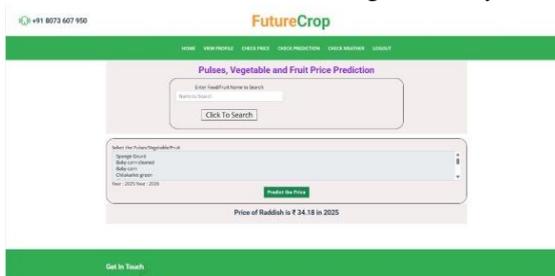


FIGURE 7 - This page allows the farmer to search and select a crop, then predict its future price using the trained machine learning model. After selection, the system displays the estimated price for the chosen year instantly on the screen.

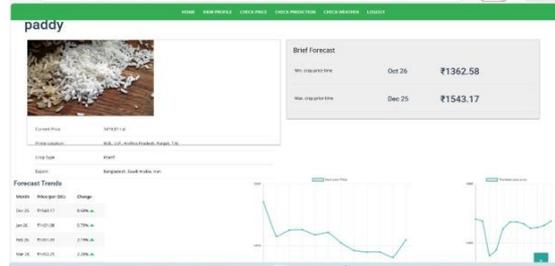


FIGURE 8 - This page provides a detailed price forecast for the selected crop, showing current price, major producing states, and export countries. It also displays monthly trends along with a one-year future price prediction using machine learning graphs and analysis.



FIGURE 9 - This page allows farmers to check real-time weather details by entering a city name. It displays temperature, humidity, cloud cover, and a brief weather description to help in crop planning.

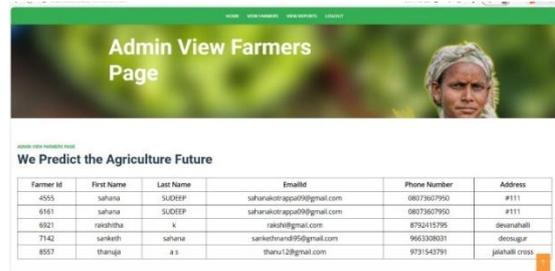


FIGURE 10 - This page allows the admin to view all registered farmers along with their details such as name, contact, and address. It helps in monitoring user activity and managing the farmer database efficiently.



FIGURE 11 - This page displays all crop price reports predicted by farmers, including the item name, predicted value, and date/time. A pie chart also shows the count of predictions crop-wise, helping admins analyze usage and trends.

VI. DISCUSSION

The results demonstrate that combining classical machine learning techniques with a lightweight web framework can effectively support agricultural decision-making. The prediction accuracy, together with the observed seasonal patterns, provides users with meaningful insights into expected price trends. While the model does not attempt to predict long-range economic shocks, it performs reliably for short-term, season-based forecasting.

Moreover, the system's modular design allows integration of additional factors such as rainfall, temperature, or real-time market feeds in future versions. This flexibility makes the system suitable for deployment in regions with diverse agricultural practices.

Overall, the proposed system shows strong potential to enhance price awareness, encourage informed crop planning, and support digital transformation in the agricultural sector.

VII. CONCLUSION AND FUTURE WORK

A. CONCLUSION

The objective of this research was to design and develop an AI-powered crop price prediction system capable of assisting farmers and agricultural stakeholders in understanding market fluctuations and making data-driven decisions. By utilizing historical vegetable and fruit price datasets and integrating them with a Decision Tree Regression model, the system proved effective in identifying non-linear price patterns and seasonal behaviors that significantly influence market dynamics.

The Decision Tree Regression algorithm consistently produced higher R-squared values compared with other baseline regression techniques, demonstrating its ability to adapt to irregular and abrupt price variations commonly seen in the agricultural sector. This made it particularly suitable for commodities such as tomatoes, onions, potatoes, and various fruits whose prices are sensitive to

climatic conditions, demand-supply imbalance, and storage limitations.

The visual exploration of data through year-wise and month-wise boxplots provided important insights into the seasonal trends inherent in agricultural pricing. The deployed web-based system, developed using Flask and supported by Firebase, ensured an efficient, interactive, and user-friendly environment for price prediction. The inclusion of optional SMS alerts through Twilio further enhanced accessibility, making the system beneficial even for users with limited internet access.

Overall, this study demonstrates the potential of integrating machine learning models with real-time web applications to support digital agriculture. The system not only offers accurate short-term price predictions but also creates valuable awareness about market volatility, enabling farmers to plan harvesting, storage, and selling strategies more effectively. The results clearly show that AI-assisted solutions can play a crucial role in stabilizing farmers' income and reducing market uncertainty.

B. FUTURE WORK

Although the proposed crop price prediction system provides reliable short-term forecasting, several research opportunities remain to further improve its accuracy, usability, and scalability. Future enhancements can be explored in the following directions

I. Integration of real-time weather and climate intelligence: Agricultural prices are highly sensitive to rainfall, temperature variations, humidity levels, and extreme climate events. Incorporating real-time meteorological data from APIs such as Open Weather or government climate portals could significantly enhance prediction quality. Coupling climatic parameters with historical pricing may allow the system to detect early-warning indicators of price surges caused by crop failures or seasonal shortages.

II. Adoption of advanced machine learning models: While Decision Tree Regression performed exceptionally well in this study, advanced ensemble-based models such as Gradient Boosting, XGBoost, and Random Forest could provide a more robust forecasting framework. These algorithms can handle noise, missing values, and multi-feature interactions

more effectively. Similarly, the inclusion of regularized regression models may help stabilize predictions for extremely volatile crops.

III. Time series forecasting for long term planning: The current system focuses on short-term predictions based on historical price patterns. Incorporating deep learning models such as LSTM, GRU, CNN-LSTM hybrids, and Facebook Prophet can enable long-term forecasting on a weekly, monthly, or annual basis. This would help farmers anticipate price cycles and plan crop cultivation, storage, or distribution strategies more effectively.

IV. Geographical market intelligence and spatial analysis: Future enhancements could incorporate geographical variations in pricing. Integrating GIS-based mapping, soil datasets, and state-wise market data would allow region-specific predictions. This would provide more localized insights, enabling farmers from different districts to receive personalized recommendations relevant to their environment.

V. Market shock Detection and anomaly prediction: Sudden events such as transport strikes, pest outbreaks, floods, or policy changes often cause abrupt price fluctuations. Future versions of the system may incorporate anomaly detection algorithms to identify such unusual behaviours. Early alerts could assist farmers and traders in preparing for market disruptions.

VI. Block-chain based price transparency: In future work, blockchain can be incorporated to ensure transparency and integrity in price records. This would prevent manipulation of market prices and enhance trust among farmers and traders.

ACKNOWLEDGEMENT

We extend our heartfelt thanks to our project guide, Mrs. Dhivya V, Assistant Professor, Department of Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology, for her continuous guidance, constructive feedback, and constant encouragement throughout the work.

We are also grateful to our Project Coordinators, Ms. Chandana K. R. and Mr. Nagendra R., Assistant Professors, Department of CSE, Sir MVIT, for their support, timely suggestions, and motivation during

every phase of the project. Our sincere appreciation also goes to Dr. Anitha T. N., Head of the Department, CSE, Sir MVIT, for providing us with a supportive academic environment.

Lastly, we thank the Department of Computer Science and Engineering and Sir M. Visvesvaraya Institute of Technology for offering the essential infrastructure, technical resources, and facilities that enabled us to successfully complete this work.

REFERENCES

- [1] G. G. Ramesh, S. S. Shinde, and A. P. Deshmukh, "Agricultural price prediction using machine learning techniques," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 45–52, 2021.
- [2] A. Singh and R. Kaur, "Crop price forecasting using decision tree and random forest models," *International Journal of Computer Applications*, vol. 178, no. 40, pp. 1–6, 2019.
- [3] K. V. Reddy and P. V. Rao, "Machine learning approaches for agricultural market analysis: A survey," *Journal of Applied Data Science*, vol. 4, no. 2, pp. 78–90, 2022.
- [4] S. Patel, M. Shah, and D. Chauhan, "Time-series forecasting for agricultural commodities using regression and tree-based models," *Procedia Computer Science*, vol. 196, pp. 538–547, 2022.
- [5] Government of India, Ministry of Agriculture and Farmers' Welfare, "Agri-market price data," National Agriculture Market (eNAM), 2023. [Online]. Available: <https://agmarknet.gov.in>
- [6] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 4th ed. San Francisco, CA, USA: Morgan Kaufmann, 2021.
- [7] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed., O'Reilly Media, 2022.
- [8] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, CA, USA, 1984.
- [9] Google, "Firebase real-time database documentation," 2023. [Online]. Available: <https://firebase.google.com/docs>
- [10] Twilio Inc., "SMS API documentation," 2023. [Online]. Available: <https://www.twilio.com/docs/sms>

[11]F. Chollet, *Deep Learning with Python*, 2nd ed.
Manning Publications, 2021.