

BUSINESS ORIENTED FOOD PRICE FORECASTING USING MACHINE LEARNING

J. Abhishek¹, N. Seavanth², B. Rohith³, CH. Varun kumar⁴,
Mrs. N.V.L. Manaswini⁵

^{1,2,3,4}Student, Department of Computer Science & Engineering

Andhra Loyola Institute of Engineering and Technology, ITI Road, Polytechnic Post, Vijayawada,
Andhra Pradesh, India

⁵Assistant Professor, Department of Computer Science and Engineering

Andhra Loyola Institute of Engineering and Technology, ITI Road, Polytechnic Post, Vijayawada,
Andhra Pradesh, India

Email id: abhishekjajimoggala@gmail.com, seavanthnelluri609@gmail.com,
bathularohithkumar30@gmail.com, chvarun987gmail.com

Abstract: Food price prediction is an important task in the agricultural sector because market prices change frequently due to various factors such as weather conditions, demand and supply, transportation costs, and seasonal variations. Farmers and traders often face difficulties in deciding the best time to sell their products because they do not have accurate information about future market prices. To address this problem, this project proposes a **Food Price Prediction System** that uses machine learning techniques to analyze historical market data and forecast future prices of agricultural commodities. The system mainly focuses on predicting prices based on factors such as commodity type, state, month, year, and weather conditions.

In this project, the **Random Forest algorithm** is used to build the prediction model because it provides better accuracy and handles large datasets efficiently. The system collects historical price data and weather information, preprocesses the dataset, and trains the machine learning model to identify patterns in price fluctuations. After training, the model can predict future food prices and display the results through a user-friendly dashboard using charts and graphs. The proposed system helps farmers, traders, and policymakers make better decisions by providing reliable price predictions, thereby reducing financial risks and improving agricultural market planning.

Keywords: Food Price Prediction, Machine Learning, Random Forest Algorithm, Agriculture Data Analysis, Commodity Price Forecasting, Weather Data Analysis, Predictive Analytics.

1. INTRODUCTION

Agriculture is a vital sector of the economy and plays a major role in ensuring food security for a country. However, the agricultural market is highly unpredictable due to various factors such as seasonal changes, climate variations, supply and demand imbalance, and government policies. These factors directly affect the prices of agricultural commodities like tomatoes, onions, potatoes, and other vegetables. Price fluctuations create uncertainty for farmers, traders, and consumers, and often lead to financial losses. Farmers may sell their crops at low prices due to fear of price drops, while traders may face losses due to sudden price decreases or unexpected market demand changes. Therefore, accurate food price forecasting is essential to help stakeholders make informed business decisions.

Traditional methods of price forecasting, such as basic statistical models, are limited in their ability to handle complex and non-linear patterns in market data. Agricultural price data is affected by multiple variables and seasonal trends, which makes it difficult for simple models to provide accurate predictions. Machine learning offers a more effective solution because it can analyze large datasets and learn hidden patterns in data. Machine learning models are capable of capturing complex relationships between input features and output values, making them suitable for predicting future prices based on historical data.

In this project, a business-oriented food price forecasting system is developed using machine learning techniques. The main aim of the project is to provide accurate predictions of future food prices to support better business planning. The system focuses on agricultural commodities such as tomatoes and other vegetables that are frequently traded in local markets. By predicting future prices, the system helps farmers decide the best time to sell their crops, traders to plan their inventory, and retailers to set better pricing

strategies. This approach improves the overall efficiency of the agricultural supply chain and reduces economic losses caused by price uncertainty.

The project uses two major machine learning models: Random Forest and Long Short-Term Memory (LSTM). Random Forest is a robust machine learning algorithm that works well with structured data and can handle non-linear relationships. It uses multiple decision trees to improve prediction accuracy and reduce overfitting. LSTM is a deep learning model specifically designed for time-series forecasting. It can learn long-term dependencies and seasonal patterns in sequential data. Both models are trained using historical price data collected from agricultural market sources. The dataset includes information such as date, commodity price, market location, and other relevant features. Data preprocessing techniques such as handling missing values, outlier detection, and normalization are applied to improve the quality of the dataset.

The system is evaluated using performance metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics help measure the accuracy of the predictions and compare the performance of different models. The experimental results show that machine learning models, especially LSTM, provide better accuracy compared to traditional forecasting methods. The project also demonstrates the importance of combining data science with business-oriented thinking to develop practical solutions for real-world problems.

2. Literature Survey

The literature on food price prediction using economic indicators and machine learning techniques has expanded significantly in recent years. The reviewed studies can be categorized under two headings: relationship and causality and prediction models.

2018-Soltani

In the year 2018 category of Prediction Model and Methodology Approach is ANFIS (Adaptive Neuro-Fuzzy Inference System) combining fuzzy logic and neural networks. In this content is Soltani and Pooya design an AI system to predict the success of new food products. The ANFIS algorithm integrates fuzzy logic and neural networks, processing data from diverse sources such as market research and social media to forecast product performance.

2020-Pooya

In the year 2020 category of Prediction Model and methodologies approach is Applicability of ANN, SVM, genetic algorithms, and hybrid techniques in stock price forecasting. In this content we have conduct a study on stock price forecasting. Their findings suggest that: Artificial Neural Networks (ANN) are best suited for predicting numerical stock index values; Support Vector Machines (SVM) perform well in classification tasks, such as predicting market direction; Hybrid ML techniques may overcome limitations of single-method approaches.

2019- Selim and Demir Kiran

In the year 2019 category of Prediction Model and Methodology Approach is Superiority of ANN over logarithmic regression. In this content we have Selim and Demir Kiran analyze household budget survey data from TÜİK to identify factors affecting food expenditures and track temporal changes. They develop predictive models using logarithmic regression and Artificial Neural Networks (ANN). Results show that the ANN model outperforms the semi-logarithmic regression model in forecasting accuracy.

2021- Sarangi

In the year 2021 category of Prediction Model and Methodology Approach is Backpropagation-trained ANN model for CPI forecasting Evaluation: MAPE, and the content of this Sarangi aim to forecast the Consumer Food Price Index (CFPI) in India using a machine learning approach. A backpropagation-trained Artificial Neural Network (ANN) is implemented using the Zaitun statistical software. MAPE values are

used to validate model accuracy, which is reported to be very high, indicating strong predictive performance.

2023- Fan explore

In the year 2023 category of Relationship and Causality, Methodology Approach is Panel Structural Vector Autoregression (PSVAR) technique to assess inflation effects. The content of Fan explore the relationship between information asset investments and inflation. Utilizing the PSVAR method, they analyse both short- and long-term dynamics. Their findings suggest that low to moderate inflation levels are positively correlated with the market value of R&D firms, whereas high inflation has a negative effect.

3. Proposed System

The proposed methodology follows a structured and modular approach to ensure flexibility and ease of implementation. The framework consists of data collection, preprocessing, model training, forecasting, and decision support stages. Data is collected from reliable public sources such as commodity price databases and market reports. This ensures transparency and reproducibility.

Preprocessing steps include handling missing values, normalization, and trend analysis. These steps improve data quality and model performance. Seasonality is an important characteristic of food prices. Seasonal patterns are identified and incorporated into the modeling process. Machine learning models are selected based on their ability to capture nonlinear price behavior while remaining computationally efficient.

Hybrid approaches that combine statistical and machine learning methods are used to improve stability and accuracy. The training process is designed to be fast and resource-efficient, allowing frequent model updates as new data becomes available.

Model performance is evaluated using standard error metrics to ensure reliability. The forecasting module generates short-term and medium-term price predictions. These predictions are then passed to the decision-support component for business interpretation.

The proposed system is designed with a modular and scalable architecture. It consists of data collection, preprocessing, model training, forecasting, and decision-support components. The system relies on publicly available historical price data, ensuring transparency and ease of implementation. The architecture allows businesses to adapt the system to different food commodities and market conditions without extensive modifications.

This data is then processed using data preprocessing techniques like cleaning missing values, normalization, and trend analysis to improve data quality. After preprocessing, machine learning models are trained to identify patterns and relationships in the data. The trained model forecasts short-term and medium-term food prices, helping farmers, traders, and businesses make better decisions regarding pricing, procurement, and inventory management. The system also includes a visualization dashboard that displays price trends and prediction results in charts and graphs. The proposed system is designed to be simple, scalable, cost-effective, and suitable for real-world business applications.

4. Methodology

1. Data Collection

The first step in the proposed system is **data collection**, where historical crop price data and weather-related information are gathered from reliable sources. The dataset contains details such as crop name, state, month, year, and previous market prices. In addition to historical data, the system also collects **real-time weather data** including temperature, humidity, and rainfall through a weather API. This combination of agricultural and environmental data provides a strong foundation for building an accurate crop price prediction model.

2. Data Preprocessing

After collecting the raw data, the next step is **data preprocessing**. In this phase, the collected data is cleaned and prepared for machine learning analysis. Missing values, inconsistent entries, and duplicate records are identified and removed to improve data quality. The categorical features such as crop names and states are

encoded into numerical formats so that the machine learning algorithm can process them effectively. The dataset is also organized into input features and output variables, where the input includes crop details and weather conditions, while the output represents the crop price.

3. Feature Selection

Feature selection plays an important role in improving the performance of the prediction model. In this step, the most relevant attributes that influence crop prices are identified. Important features include crop type, state, month, year, previous price, temperature, humidity, and rainfall. By selecting only the meaningful features, the system reduces computational complexity and improves prediction accuracy. These features are then used as input variables for training the machine learning model.

4. Model Training Using Machine Learning

The system uses a **machine learning algorithm such as Random Forest** to train the prediction model. Random Forest is an ensemble learning technique that builds multiple decision trees and combines their results to produce a more accurate prediction. During the training phase, the prepared dataset is divided into training and testing sets. The model learns patterns and relationships between crop prices and the selected features from the training data. This trained model is then capable of predicting future crop prices based on new input values.

5. Prediction Process

Once the model is trained, it is integrated into the backend server of the application. When a user enters crop details such as state, month, and year through the frontend interface, the system sends the request to the backend. The backend retrieves the required weather information through the API and combines it with the user input. This data is then passed to the trained machine learning model, which generates the predicted crop price along with accuracy and trend information. The prediction result is then sent back to the user interface.

6. System Integration

The entire system is integrated using a **web-based architecture**. The frontend interface is developed using HTML, CSS, and JavaScript, while the backend is implemented using Python and the Flask framework. The backend communicates with the machine learning model to perform predictions and interacts with the database to store user data and prediction results. A cloud-based database stores information such as users, predictions, feedback, and login records. This integration ensures smooth communication between all components of the system.

7. Result Visualization and Admin Monitoring

To enhance usability, the system provides graphical visualization of predictions and system usage. The admin dashboard allows administrators to monitor system performance, view prediction statistics, analyze monthly usage patterns, and review user feedback. These insights help administrators improve the system and ensure reliable service for users. Visualization also helps farmers and traders better understand price trends and make informed decisions.

5. Proposed System Hardware Results

5.1. System Interface and Working

The developed system, *AgriForecast – Agricultural Price Forecasting Portal*, provides a user-friendly interface for predicting and analyzing agricultural commodity prices. The system is designed with multiple modules such as homepage, registration, login, browsing, prediction, and admin dashboard.

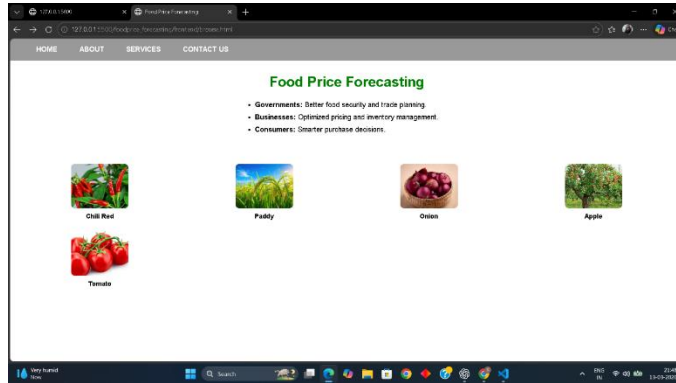


Fig:5.4 Browse Page

The browse page displays different agricultural commodities such as:

- Chili
- Paddy
- Onion
- Apple
- Tomato

Users can select a commodity of interest to view its price trends and predictions. The page also provides key benefits of the system for:

- Governments (policy planning)
- Businesses (inventory management)
- Consumers (better purchasing decisions)

5.5 Prediction of Item-1

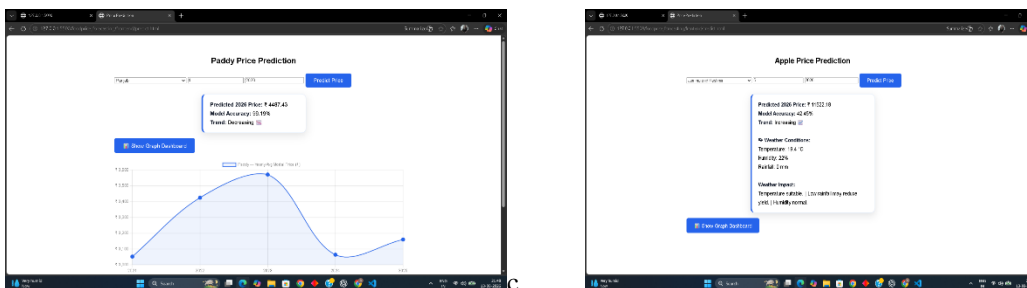


Fig:5.5 Prediction of Item-1 & Testcase-1

This module displays detailed price predictions for a selected commodity (e.g., Onion). The system provides:

- State-wise data
- Monthly selection options
- Average prices for previous years (2024, 2025)
- Predicted prices for future years (2026)

Users can filter data based on commodity and month, making the system highly interactive and analytical.

5.6 Admin Dashboard

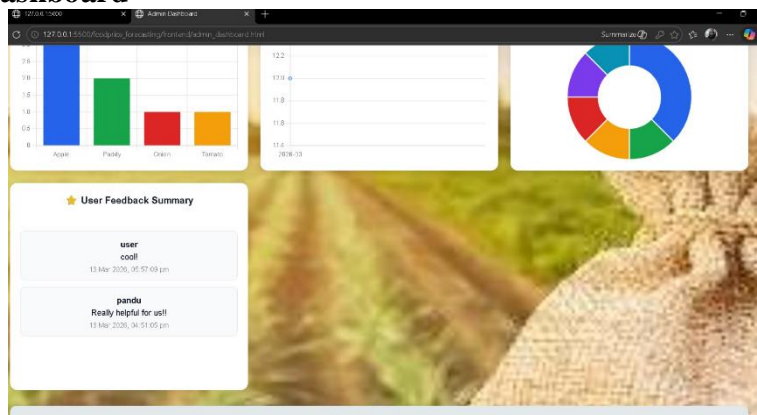


Fig:5.6 Admin Page

The admin dashboard allows administrators to manage system data and monitor activities. It includes features such as:

- Data management
- User management
- Updating commodity prices
- Monitoring predictions

This ensures smooth functioning and maintenance of the system.

6. CONCLUSION

The **Business-Oriented Food Price Forecasting using Machine Learning** project focuses on predicting future food prices by analyzing historical market data. The system uses machine learning techniques to identify patterns and trends in food price fluctuations.

By applying data preprocessing, model training, and forecasting methods, the system can generate accurate price predictions for different food commodities. This helps farmers, traders, and businesses make better decisions about selling, buying, and managing inventory. The proposed system is simple, scalable, and cost-effective because it uses publicly available data and efficient machine learning models. Overall, the project demonstrates how machine learning can improve price prediction and support better planning in agricultural and food markets.

REFERENCES

1. Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly Media, 2019.
2. Sebastian Raschka and Vahid Mirjalili, *Python Machine Learning*, Packt Publishing, 2017.
3. Andreas C. Müller and Sarah Guido, *Introduction to Machine Learning with Python*, O'Reilly Media, 2016.
4. Food and Agriculture Organization, *Food Price Index Reports*, FAO Publications.
5. World Bank, *Commodity Market Outlook Reports*, World Bank Publications.
6. Institute of Electrical and Electronics Engineers, *Research Papers on Food Price Forecasting and Machine Learning Applications*.
7. Elsevier Journals, *Articles on Agricultural Market Analysis and Commodity Price Prediction*.
8. International Monetary Fund, *Global Economic Outlook Reports*, IMF Publications.