AI Based Model for Predicting Prices of Agri-Horticultural Commodities

Mr. Vijayakumar M
Assistant Professor
B.Tech Artificial Intelligence and Data Science
Dr. Mahalingam College of Engineering and Technology
Coimbatore, India
vijayakumarm@drmcet.ac.in

Sharanya T
B.Tech Artificial Intelligence and Data Science
Dr. Mahalingam College of Engineering and Technology
Coimbatore, India
sha2005ranya@gmail.com

Santhosh S
B.Tech Artificial Intelligence and Data Science
Dr. Mahalingam College of Engineering and Technology
Coimbatore, India
santhoshsanthosh24133@gmail.com

Aashif Shadin K N
B.Tech Artificial Intelligence and Data Science
Dr. Mahalingam College of Engineering and Technology
Coimbatore, India
aashifanshaf786@gmail.com

Abstract -- Variations in agricultural commodity prices pose major challenges for farmers, traders, and policymakers, impacting economic stability and food security. Traditional models for price prediction often lack the accuracy and flexibility required for effective decision-making. This study introduces a data-driven framework that leverages advanced machine learning techniques to enhance the reliability of agricultural price forecasts. The framework integrates multiple data sources, including real-time market trends, historical price patterns, weather conditions, supply-demand fluctuations, and government policies. Models such as ARIMA, SARIMA are applied to identify complex patterns in price movements This approach not only improves forecast accuracy but also supports better decision-making in buffer stock management, procurement strategies, and policy interventions, contributing to a more stable agricultural ecosystem.

***Keywords*—Agricultural price prediction, Machine learning, ARIMA, SARIMA, Time Series Forecasting, Market analysis, Economic forecasting, Decision support.**

I. INTRODUCTION

Agricultural commodity price fluctuations significantly affect the livelihoods of farmers, influence market stability, and pose challenges for policymakers in ensuring food security. Price volatility arises due to multiple factors, including unpredictable weather patterns, supply-demand imbalances, changes in government policies, and global economic conditions. Inaccurate price predictions lead to poor decision-making, impacting procurement strategies, buffer stock management, and overall market interventions.

Traditional price prediction models, such as linear regression and basic time series models, often fail to capture the complex, non-linear relationships that drive commodity price changes. As a result, stakeholders face difficulties in making informed decisions. To address these limitations, this research proposes a robust, data-driven framework that employs advanced machine learning (ML) techniques to predict agricultural prices with higher accuracy. The framework utilizes diverse data sources such as real-time market data, historical price trends, weather information, and economic indicators to build comprehensive predictive models.

By combining predictive accuracy with model interpretability, this approach enhances strategic decision-making in agricultural supply chains. It aids in planning buffer stock management, optimizing procurement, and formulating effective market interventions. This research contributes to the advancement of agricultural price forecasting by integrating state-of-the-art machine learning techniques with explainability, fostering a more resilient and data-driven agricultural ecosystem.

II. RELATED WORKS

Saman Ghaffarian, Firouzeh Rosa Taghikhah, and Holger R. Maier [1] investigate the role of machine learning models, including decision trees and ensemble methods, in disaster prediction. Their findings highlight the use of Random Forest (RF) and XGBoost as commonly employed techniques for agricultural market predictions. By leveraging decision tree-based models, these approaches provide clear interpretability, which is essential for understanding the relationships between agricultural prices and influencing factors such as weather and government policies. The study emphasizes how XGBoost helps improve prediction accuracy while maintaining the interpretability necessary for decision-makers in the agricultural sector.

Bukhoree Sahoh and Anant Choksuriwong [2] discuss the limitations of traditional AI systems in critical decision-making scenarios, especially within the agricultural sector. They note that while AI systems excel in processing large amounts of complex data, the lack of transparency in many models hinders decision-makers from fully trusting the predictions. They propose using Explainable AI (XAI) to offer interpretable predictions, making it easier to understand which factors, such as crop yields or global economic indicators, contribute to price fluctuations.

The use of Explainable AI (XAI) in the context of agricultural commodity price prediction has gained significant attention due to its ability to improve model transparency and trust, which is crucial in decision-making for farmers, traders, and policymakers. Several studies have explored the integration of AI and XAI techniques to predict agricultural commodity prices more accurately and explain the reasoning behind predictions.

This research underscores the importance of transparent AI systems in agricultural forecasting to guide strategic decisions like pricing and procurement.

T. Sri Sai Charan, U. Rohit Reddy, T. Samara Simha Reddy, and Tarun G [3] focus on the application of machine learning (ML) and deep learning (DL) techniques for predicting agricultural commodity prices in India, particularly in flood-prone regions where weather plays a significant role in price volatility. The study highlights the importance of incorporating external factors like weather data, market policies, and supply-demand changes into forecasting models**.** LSTM (Long Short-Term Memory) networks are identified as particularly useful for capturing long-term dependencies in price movements, but the authors emphasize the need for XAI methods to explain how these external factors influence the predictions, which aids in better policy intervention and market planning.

Kumar et al. [5] review various AI and machine learning techniques applied to agricultural commodity price forecasting, with a particular focus on the integration of XAI. They highlight how ARIMA, SARIMA, and XGBoost have been used to predict price trends but stress the growing importance of incorporating explainability into these models. The study emphasizes that the use of XAI tools like SHAP can help stakeholders identify and understand the economic and environmental factors that drive price fluctuations, providing them with actionable insights to make informed decisions about market interventions and supply chain management.

Yuan et al. [6] propose an innovative hybrid model combining LSTM and ARIMA for agricultural commodity price prediction. They integrate SHAP to interpret the model's decision-making process, enabling stakeholders to understand how various factors, such as rainfall, supply disruptions, and government policies, influence the final price predictions. Their approach demonstrates that integrating traditional time-series methods with advanced neural networks and explainable models can significantly improve forecasting performance while providing the transparency needed for stakeholders to trust and act upon the predictions.

III. PROBLEM DESCRIPTION

The Department of Consumer Affairs monitors the prices of 22 essential food commodities through 550 price reporting centers nationwide. It also manages buffer stocks for commodities like pulses (gram, tur, urad, moon, and masur) and onions, which are strategically released into the market to mitigate price volatility. Market interventions, such as the release of these stocks, are made based on the analysis of price trends, seasonal variations, and production estimates. Currently, price analysis largely relies on seasonal patterns, historical data, market intelligence, and crop forecasts. While ARIMA-based economic models are utilized to predict pulse prices, these methods often face limitations in accuracy and adaptability.

This project seeks to develop an AI-based agricultural commodity price prediction system that improves forecasting accuracy by incorporating deep learning models like SARIMA and ARIMA. By integrating real-time data, historical trends, seasonal patterns, and external factors affecting the market, this system aims to predict price fluctuations with greater precision.

The objective is to provide farmers, traders, and policymakers with valuable insights to better manage buffer stocks, ensure price stability, and mitigate risks. Ultimately, this system will contribute to reducing economic uncertainty and enhancing decision-making in the agricultural sector.

IV. OBJECTIVE OF THE PROJECT

The objective of this project is to develop an AI-driven agricultural commodity price prediction system that significantly enhances forecasting accuracy and facilitates strategic market interventions. By integrating deep learning models such as ARIMA, and SARIMA with real-time data, the system aims to analyze historical trends, seasonal patterns, and external market factors, including weather conditions, supply-demand fluctuations, and government policies. The primary goal is to provide farmers, traders, and policymakers with actionable, data-driven insights that will support better decision-making in areas such as buffer stock management, price stabilization, and risk mitigation. Ultimately, this project seeks to reduce economic uncertainty in the agricultural sector by offering a more reliable and transparent method of predicting price fluctuations, helping to stabilize markets and improve overall economic stability.

V. METHODOLOGY

a) Data Collection and Preprocessing

Dataset Selection: Historical data on agricultural commodities, including crop yield, market prices, weather patterns, and soil quality, is collected from government databases, meteorological agencies, and agricultural research centers.

Data Cleaning: Missing values are imputed, outliers removed, and data standardized.

Feature Engineering: Features such as rainfall intensity, temperature variations, soil moisture levels, and economic ***indicators are extracted.***

Data Splitting: The dataset is divided into training (70%), validation (15%), and test (15%) sets.

b) Model Selection and Development

Baseline Models: Traditional models like ARIMA and regression methods are implemented for baseline comparisons.

Hyperparameter Tuning**:** Optimization techniques such as grid search and Bayesian optimization are applied to enhance model performance.

Model Visualization: Techniques such as SHAP summary plots and feature attribution maps illustrate the decision-making process.

Fairness and Bias Evaluation: Bias assessments ensure equitable predictions, reducing potential disadvantages for small-scale farmers.

c) Performance Evaluation

*Evaluation Metrics***:** RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), R-squared, and MAPE (Mean Absolute Percentage Error) are used to assess model accuracy.

Comparison with Baseline Models: AI-based predictions are benchmarked against traditional statistical methods.

*Ablation Studies***:** Experiments are conducted to analyze the impact of different features and model architectures.

d) Deployment and Real-World Validation

Prototype Development: A web-based or mobile application is designed for real-time commodity price prediction.

Real-Time Prediction: Integration of weather APIs and market data feeds enables dynamic updates to forecasts.

User Feedback and Adaptation: Feedback from farmers, traders, and policymakers is collected to improve model usability and accuracy.

VI. PROPOSED SYSTEM

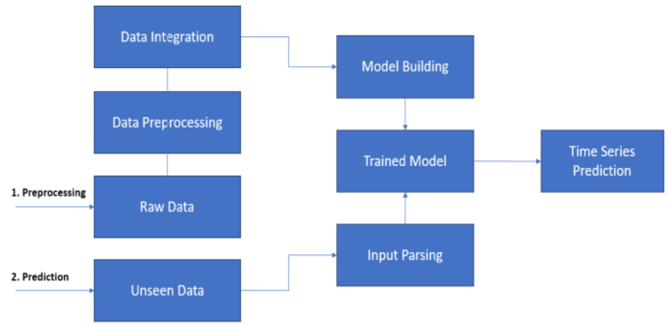

Fig 1 :Flow chart: AI based agricultural commodities prediction

1) Data Collection: The data collection process forms the foundation of the system. The system gathers historical and real-time data from the following key sources:

AGMARKNET: A government portal providing data on market prices, arrivals, and trade of agricultural commodities across India.

Government Reports: Reports and publications from agricultural ministries and organizations offering valuable historical data.

Market APIs: Real-time market price data through API integrations from trading platforms and market exchanges.

News/Social Media: Sentiment analysis on market trends and commodity prices based on news articles, reports, and social media content related to agriculture.

2) Data Preprocessing

The preprocessing stage ensures that the collected data is clean, consistent, and ready for analysis and modeling:

Data Cleaning: Missing values are imputed, and outliers are detected and managed to avoid biases in the prediction models.

Feature Engineering: Additional features are created, including economic indicators, weather patterns, and soil conditions, that might influence commodity prices.

Time-Series Processing: Data is organized and transformed to be suitable for time-series forecasting models, considering seasonal trends, cycles, and historical price movements.

3) Price Prediction Models

Various machine learning and statistical models are used for predicting agricultural commodity prices:

ARIMA: Autoregressive Integrated Moving Average models are used for capturing the temporal dependencies in price movements and for forecasting future prices based on past trends.

Random Forest: An ensemble learning technique that aggregates predictions from multiple decision trees, improving accuracy by reducing overfitting and capturing complex relationships.

LSTM and GRU: Recurrent neural networks specifically designed for time-series data. These models are capable of learning long-term dependencies in sequential data, such as commodity price trends and weather patterns.

4) Volatility Analysis

To assess the stability of commodity prices and detect potential price fluctuations, the system incorporates the following:

GARCH Model: The Generalized Autoregressive Conditional Heteroskedasticity model helps to capture and forecast volatility in commodity prices over time, predicting periods of high risk or uncertainty.

Anomaly Detection: Algorithms are used to identify irregularities or outliers in the price data, which can signal unexpected events that may affect commodity markets.

5) Decision Support

The system incorporates a decision support layer that helps stakeholders make informed choices:

Buffer Stock Optimization: Using optimization techniques to determine the ideal stock levels for agricultural commodities, minimizing market disruption during shortages and controlling price inflation.

Explainable AI: SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) are integrated to explain the predictions of the machine learning models. This ensures transparency in decision-making, allowing users to understand the reasons behind the predictions.

VII. MODULE DESCRIPTION

The system is divided into three modules:

7.1. Front-End Interface (User Interaction Console)

Framework: Streamlit (Python) or Flask/Dash (for web-based UI).

Functionality: **Commodity/Location Selection**: Dropdown menus for users to choose commodities (e.g., onion, tomato) and regional markets (e.g., Bengaluru, Coimbatore).

Visualizations: Interactive plots (Plotly/Matplotlib) showing historical price trends and forecasted values.

Results Display: Tabular data of predicted prices with confidence intervals for the next 7–30 days.

Usability:

Responsive Design: Adapts to desktop and mobile screens.

Error Handling: Validates user inputs (e.g., invalid date ranges) with clear prompts.

Export Options: Download forecasts as CSV/PDF for further analysis.

7.2. Back-End Processing (Prediction Engine)

Core Models:

ARIMA: Autoregressive Integrated Moving Average for baseline price forecasting.

SARIMA: Seasonal ARIMA to capture cyclical patterns (e.g., annual price fluctuations).

Data Pipeline:

Data Cleaning: Handles missing values, outliers, and normalizes Agmarknet datasets.

Feature Engineering: Focuses on temporal features (lagged prices, rolling averages).

Significance: Processes raw market data into actionable forecasts with measurable accuracy (MAE/RMSE).

7.3. Interactive Analytics Workflow

Functionality:

Dynamic Forecasting: Adjust forecast horizons (short-term vs. long-term) and compare results.

Model Evaluation: Displays performance metrics (MAE, RMSE) for transparency.

Scenario Testing: Simulate price impacts of external factors (e.g., monsoon delays) via hypothetical inputs.

Importance:Bridges the gap between statistical models and end-user decision-making (farmers, policymakers).

VII. SNAPSHOTS

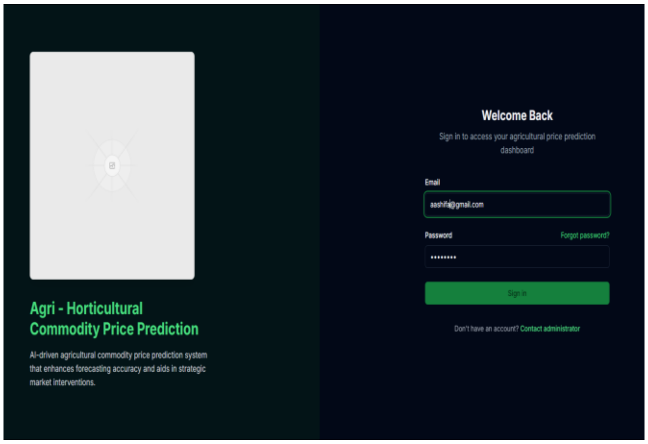

Fig:1 Login Page

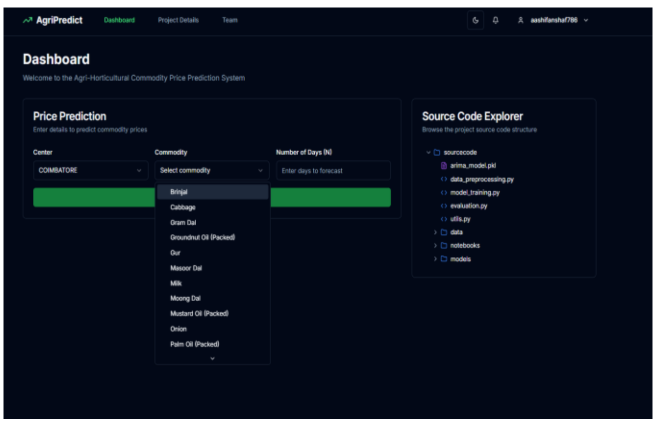

Fig:2 Dashboard

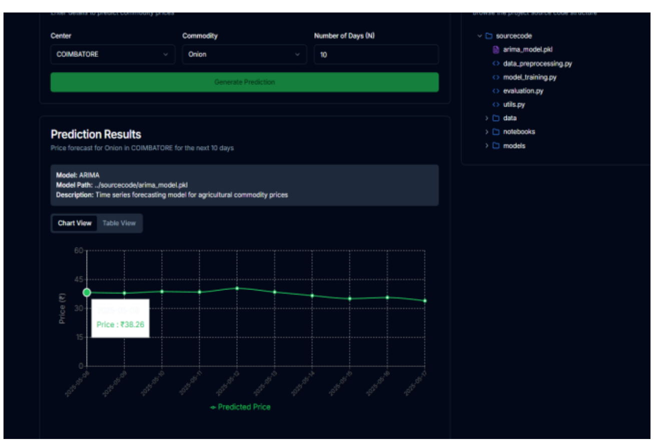

Fig:3 Price Forecasting

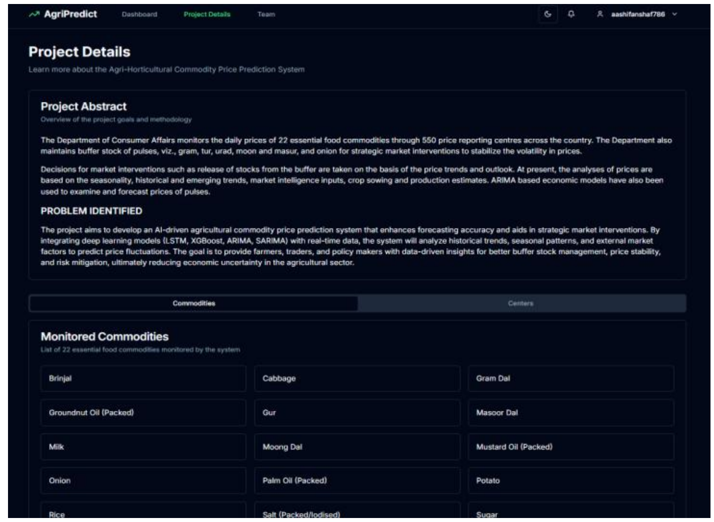

Fig: 4 Project Details

VIII. CONCLUSION AND FUTURE WORK

This project focused on forecasting agricultural commodity prices using ARIMA and SARIMA time series models, based on historical data from Agmarknet across multiple Indian cities (1997–2015). After thorough preprocessing—including filtering, handling missing values, and resampling—the models provided reliable short-term forecasts, offering valuable insights for market planning. The system's strength lies in its practical application for farmers, vendors, and policymakers, enabling better decision-making through price predictions. However, traditional models like ARIMA/SARIMA have limitations in capturing non-linear trends and adapting to sudden market shifts. Future improvements could integrate advanced models such as LSTM, GRU, or Prophet to better model complex patterns. Adding external factors like weather, inflation, and policy changes can also enhance forecast accuracy by accounting for real-world influences. Developing a user-friendly web or mobile interface would make the tool more accessible, allowing users to input parameters and receive visual forecasts. Incorporating Explainable AI (XAI) techniques would further improve transparency, crucial for policy and public use. Finally, automating the data pipeline to fetch and update real-time market data ensures the model remains current and adaptive. Together, these enhancements can significantly improve forecasting performance, scalability, and impact in agri-tech and economic planning.

IX. REFERENCES

[1] Elbasi, E., Mostafa, N., Alarnaout, Z., et al. (2022). Artificial Intelligence Technology in the Agricultural Sector: A Systematic Literature Review. IEEE.

[2] Ngoc-Bao-Van Le, Seo, Y. S., Huh, J. H. (2024). AgTech: Volatility Prediction for Agricultural Commodity Exchange Trading Applied Deep Learning. IEEE.

[3] Zhang, D., Chen, S., Ling, L., Xia, Q. (2020). Forecasting Agricultural Commodity Prices Using Model Selection Framework With Time Series Features. IEEE.

[4] Suhasini, S., & Reddy, P. (2018). ARIMA Model for Forecasting Agricultural Commodity Prices in India. International Journal for Research in Applied Science and Engineering Technology (IJRASET).

[5] Jadhav, P., & Shinde, R. (2021). Time Series Analysis of Agricultural Data Using SARIMA and LSTM. Springer.

[6] Chakraborty, S., & Bose, I. (2019). Machine Learning Techniques for Crop Yield and Price Prediction. Procedia Computer Science.

[7] Mandal, K., & Ghosh, S. (2020). Price Forecasting of Vegetables in India Using Hybrid Time Series Models. International Journal of Engineering Research & Technology (IJERT).

[8] Dey, P., & Samanta, D. (2022). Explainable AI in Agri Market Systems. Elsevier.

[9] Agmarknet Portal. https://agmarknet.gov.in/. Ministry of Agriculture and Farmers Welfare, Government of India.

[10] Department of Consumer Affairs – Price Monitoring Division. https://consumeraffairs.nic.in/. Government of India.

[11] FAO Food Price Index. https://www.fao.org/worldfoodsituation/foodpricesindex. Food and Agriculture Organization.

[12] Open Government Data (data.gov.in). Ministry of Agriculture, Government of India.

[13] Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control. Wiley.

[14] Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). Forecasting: Methods and Applications. Wiley.

[15] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice (3rd ed.). https://otexts.com/fpp3. OTexts.