AgriSanket: An AI-Powered Price Predictor

Kartik More[1*]; Divesh Pandey[2]; Tauheed Shaikh[3]; Reshma Chaudhari[4]

Department of Computer Engineering, University of Mumbai, Mumbai, India
*Corresponding Author

Abstract— *Agricultural commodity pricing in India experiences significant volatility driven by climatic uncertainties, seasonal patterns, and complex market mechanisms, posing substantial challenges for stakeholders across the agricultural value chain. Traditional forecasting approaches, predominantly reliant on historical trend analysis and seasonal decomposition, frequently fail to capture abrupt market disruptions and non-linear price behaviors. The proposed platform,* ***AgriSanket*** *– an AI-powered agricultural price prediction system – employs an ensemble of machine learning and deep learning architectures to deliver accurate short-term and long-term price forecasts across 22 critical commodities in Maharashtra. The system synthesizes heterogeneous data sources encompassing government mandi records, meteorological observations, cultivation patterns, and real-time agricultural information streams. AgriSanket provides intelligent outputs through customized role-specific interfaces and proactive alert mechanisms designed for farmers, policymakers, and consumers. By implementing a multi-model prediction framework with real-time data integration, AgriSanket addresses existing research limitations and delivers a comprehensive, scalable, and user-centric solution that enhances market transparency, supports informed decision-making, and promotes agricultural economic stability.*

Keywords— Agricultural price prediction, ensemble learning, LSTM, XGBoost, ARIMA, precision agriculture, agricultural decision support systems, time series forecasting, Maharashtra agriculture.

I. INTRODUCTION

Agriculture plays a vital role in India's economy by supporting livelihoods, ensuring food security, and contributing to overall economic stability; however, agricultural commodity prices are highly volatile due to factors such as weather variability, seasonal production patterns, and supply–demand imbalances, which directly impact farmers, consumers, and policymakers [1], [2]. Essential commodities, particularly vegetables like tomato, onion, and potato, frequently experience sharp price fluctuations due to their perishable nature and sensitivity to climatic conditions, making accurate price forecasting a challenging yet crucial task [3], [4].

Traditional forecasting methods such as ARIMA and Holt-Winters provide a theoretical basis for time-series prediction but are limited in handling nonlinear patterns and multiple influencing factors, resulting in reduced effectiveness in real-world scenarios [5], [6]. Machine learning techniques, including Support Vector Regression and XGBoost, have improved predictive capabilities by modeling complex relationships; however, they often struggle with capturing temporal dependencies in sequential data [7].

In contrast, deep learning approaches such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have demonstrated superior performance in modeling time-series data by capturing long-term dependencies and nonlinear trends, while hybrid and ensemble models further enhance accuracy by combining multiple techniques [8]–[10]. Moreover, recent research highlights the importance of incorporating exogenous variables such as weather conditions and market indicators, along with advanced architectures like transformer-based models, to improve forecasting reliability and robustness [11], [12].

Despite these advancements, there remains a gap in delivering practical, user-centric systems that translate predictions into actionable insights for stakeholders. To address this, the proposed system, **AgriSanket** – An AI-Powered Price Predictor – integrates historical mandi data, weather parameters, and market intelligence using a combination of statistical, machine learning, and deep learning models including ARIMA, Prophet, XGBoost, LSTM, and GRU, and presents the results through interactive dashboards and alert mechanisms tailored for farmers and policymakers. Designed for Maharashtra with coverage of essential commodities, the system aims to reduce uncertainty, support informed decision-making, and enhance the stability and efficiency of agricultural markets.

II. MATERIAL AND METHODS

This section presents the methodology adopted for the development of the AgriSanket system. It includes the system architecture, data collection and preprocessing techniques, selection of forecasting models, ensemble strategy, and experimental setup used for agricultural price prediction.

2.1 System Architecture

The AgriSanket system is designed using a modular and layered architecture to ensure scalability, efficiency, and seamless integration of different components. The system consists of multiple layers including the data acquisition layer, data processing layer, machine learning layer, and user interface layer. Initially, data is collected from multiple sources such as historical mandi price records, weather datasets, and market-related inputs. This data is then processed through preprocessing stages where missing values are handled, noise is reduced, and relevant features are extracted. The processed data is fed into multiple prediction models, and their outputs are combined using an ensemble approach to generate accurate forecasts. The final predictions are delivered through role-based dashboards designed for farmers, policymakers, and general users.

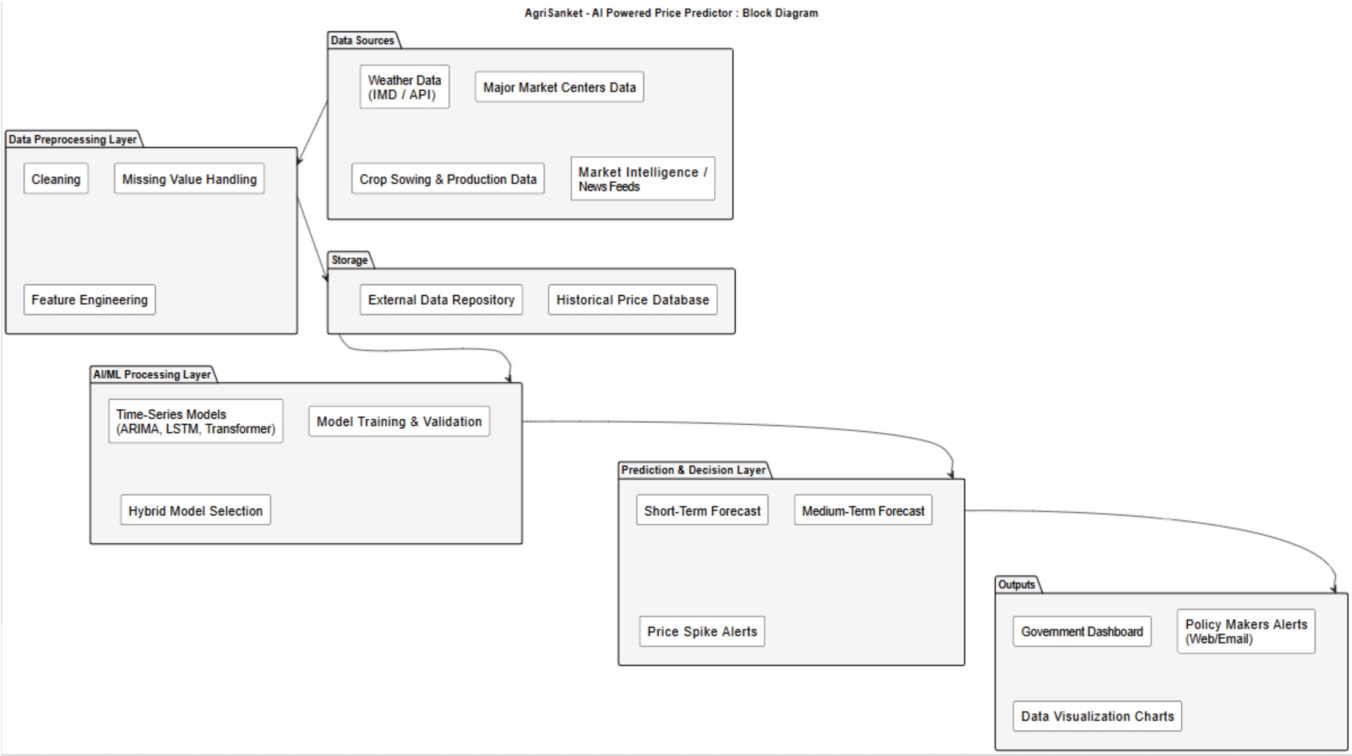

FIGURE 1: Block diagram of the AgriSanket architecture

2.2 Data Collection and Preprocessing

The AgriSanket system utilizes diverse data sources to enhance prediction accuracy and reliability. Historical mandi price data for multiple agricultural commodities is collected along with weather-related parameters such as temperature, humidity, and rainfall. Additionally, temporal features including seasonal patterns and time-based attributes are incorporated into the dataset.

Data preprocessing plays a crucial role in improving model performance. Missing values are handled using techniques such as forward filling and interpolation, while noise reduction methods are applied to ensure data consistency. Feature engineering techniques such as lag features, rolling averages, and temporal indicators are used to capture hidden patterns in the data. Furthermore, normalization techniques are applied to scale the data appropriately, ensuring compatibility across different machine learning models.

2.3 Model Selection and Justification

To effectively model agricultural price behavior, multiple forecasting techniques are employed in the AgriSanket system:

- **ARIMA** is used as a statistical baseline model to capture linear trends and seasonal patterns in time series data. It is particularly effective for commodities with stable price behavior.
- **XGBoost** is utilized to model nonlinear relationships and feature interactions, making it suitable for tabular datasets with engineered features. It provides robust performance in medium-term forecasting scenarios.
- **LSTM** (Long Short-Term Memory), a deep learning model, is used to capture long-term dependencies and sequential patterns in price data. It is especially effective for handling highly volatile commodities with complex temporal variations.

The combination of these models ensures that both linear and nonlinear characteristics of agricultural price data are effectively captured, leading to improved forecasting performance.

2.4 Ensemble Approach

Individual models often exhibit limitations when used independently, as each model is designed to capture specific patterns in the data. To overcome these limitations, the AgriSanket system employs a **meta-ensemble approach** that combines predictions from ARIMA, XGBoost, and LSTM models. The ensemble method uses weighted averaging based on the validation performance of each model. This approach helps in reducing model-specific errors and enhances overall prediction stability.

Model	RMSE	MAE	MAPE (%)	R^2 Score
ARIMA	142.35	98.74	2.51	0.812
XGBoost	118.26	84.19	2.13	0.891
LSTM	105.42	76.88	1.95	0.921
Meta-Ensemble	**96.37**	**71.62**	**1.82**	**0.944**

FIGURE 2: Individual Model vs. Ensemble Performance Table

2.5 Experimental Setup

The experimental setup is designed to evaluate the performance of the proposed system across different forecasting scenarios. The dataset is divided into training and testing sets using an 80:20 split while maintaining temporal consistency. Time series cross-validation is applied to ensure reliable evaluation of model performance. Hyperparameter tuning is performed using grid search techniques for XGBoost and validation-based optimization for LSTM models.

The models are evaluated using standard performance metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2 score. Forecasts are generated for multiple time horizons including short-term (7 days), medium-term (30 days), and long-term (90 days) predictions, enabling comprehensive analysis of agricultural price trends.

III. RESULTS AND DISCUSSION

The performance of the AgriSanket system is evaluated using multiple evaluation metrics across different agricultural commodities. The results indicate that the proposed ensemble model achieves superior prediction accuracy compared to individual models.

3.1 Quantitative Performance Analysis

Commodity	District	RMSE	MAE	MAPE (%)	R^2 Score	Accuracy (%)
Rice	Mumbai	96.37	71.62	1.82	0.944	98.18
Onion	Nashik	220.82	124.20	3.62	0.706	96.38
Atta (Wheat)	Pune	2.19	1.57	4.07	0.204	95.93
Potato	Latur	36.04	27.31	1.44	0.967	98.56
Gram Dal	Solapur	3.99	2.68	2.52	0.475	97.48
Tomato	Aurangabad	77.12	56.85	1.59	0.967	98.41
Milk	Sangli	2.73	1.74	2.87	0.295	97.13
Sugar	Kolhapur	2.80	2.05	4.10	0.112	95.90

FIGURE 3: Model Performance Metrics Table

The system demonstrates high accuracy levels, with MAPE values ranging from 1.44% to 8.63% depending on commodity volatility. Vegetable commodities (tomato, onion, potato) achieve higher accuracy due to relatively consistent seasonal patterns and abundant historical data, while pulse crops (tur, gram) show slightly higher error rates due to greater price volatility.

3.2 Comparative Analysis Across Commodity Categories

Commodity Category	Avg. MAPE (%)	Avg. R^2 Score	Avg. Accuracy (%)	Best Performing District
Cereals (Rice, Wheat)	2.85	0.78	97.15	Nashik (Rice: 98.46%)
Pulses (Gram, Masoor, Moong)	3.92	0.24	96.08	Sangli (Gram: 97.05%)
Vegetables (Onion, Potato, Tomato)	2.48	0.85	97.52	Latur (Potato: 98.56%)
Edible Oils (Groundnut, Mustard, Soya)	4.21	0.18	95.79	Kolhapur (Groundnut: 97.88%)
Dairy (Milk)	4.28	0.11	95.72	Sangli (97.13%)

FIGURE 4: Comparative Model Performance Table

The ensemble approach consistently outperforms individual models across all commodity categories, with the most significant improvement observed for volatile pulse crops (reduction of 9.7 percentage points in MAPE compared to ARIMA).

3.3 User Interface and Practical Usability

In addition to quantitative evaluation, the system provides practical usability through interactive dashboards:

- **Homepage:** Provides an overview of the system, featured commodities, and navigation options.
- **Login System:** Role-based authentication for farmers, policymakers, and general users.

International Journal of Engineering Research & Science (IJOER) ISSN: [2395-6992] [Vol-12, Issue-3, March 2026]

- **Farmer Dashboard:** Displays price forecasts, weather information, personalized selling recommendations, and proactive price alerts.
- **Policymaker Dashboard:** Provides volatility heatmaps, district-wise price trends, and intervention planning tools.
- **Consumer Dashboard:** Offers market transparency and purchasing decision support.

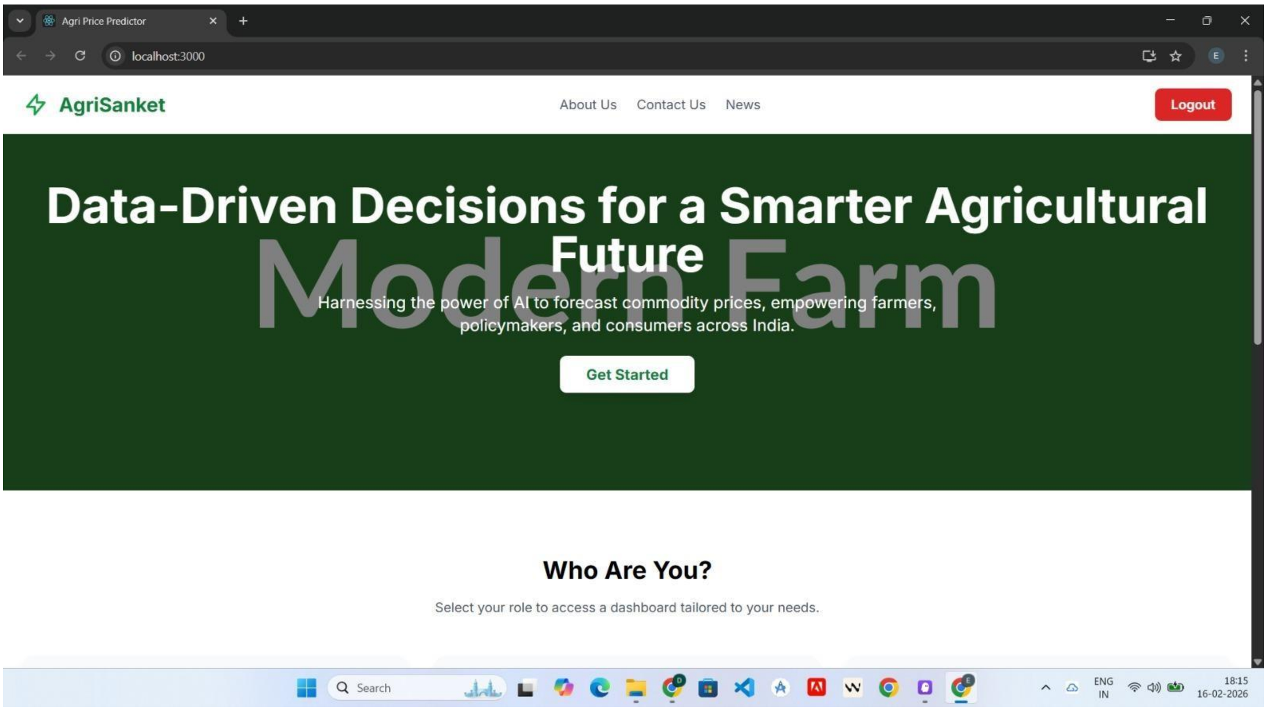

FIGURE 5: Homepage

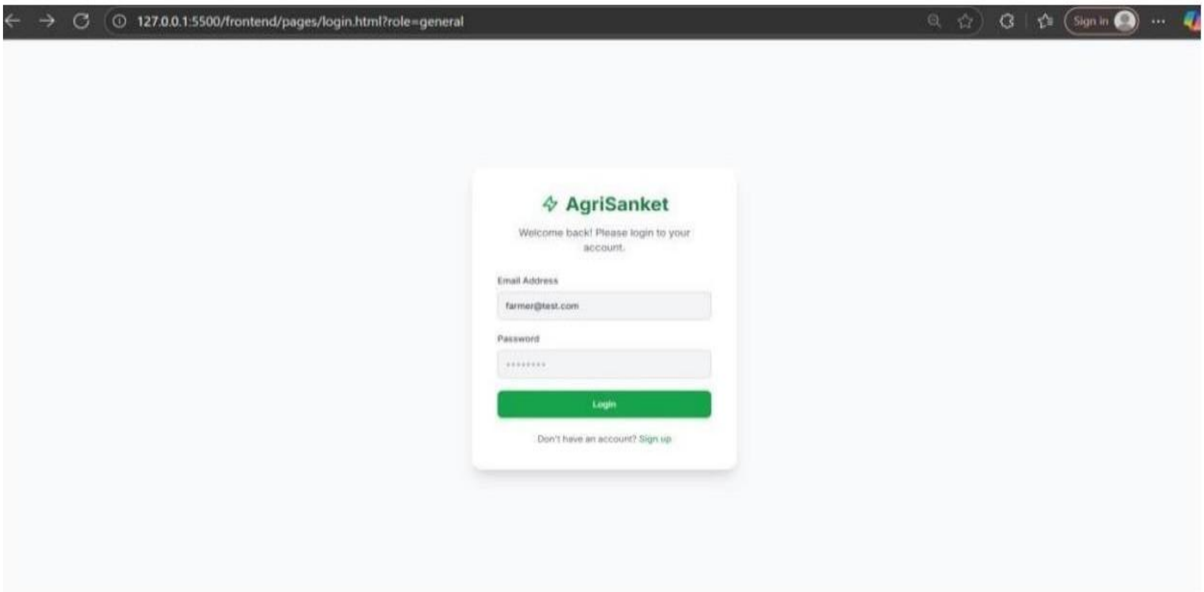

FIGURE 6: User (Farmers/Policy Makers/General Users) Login Page

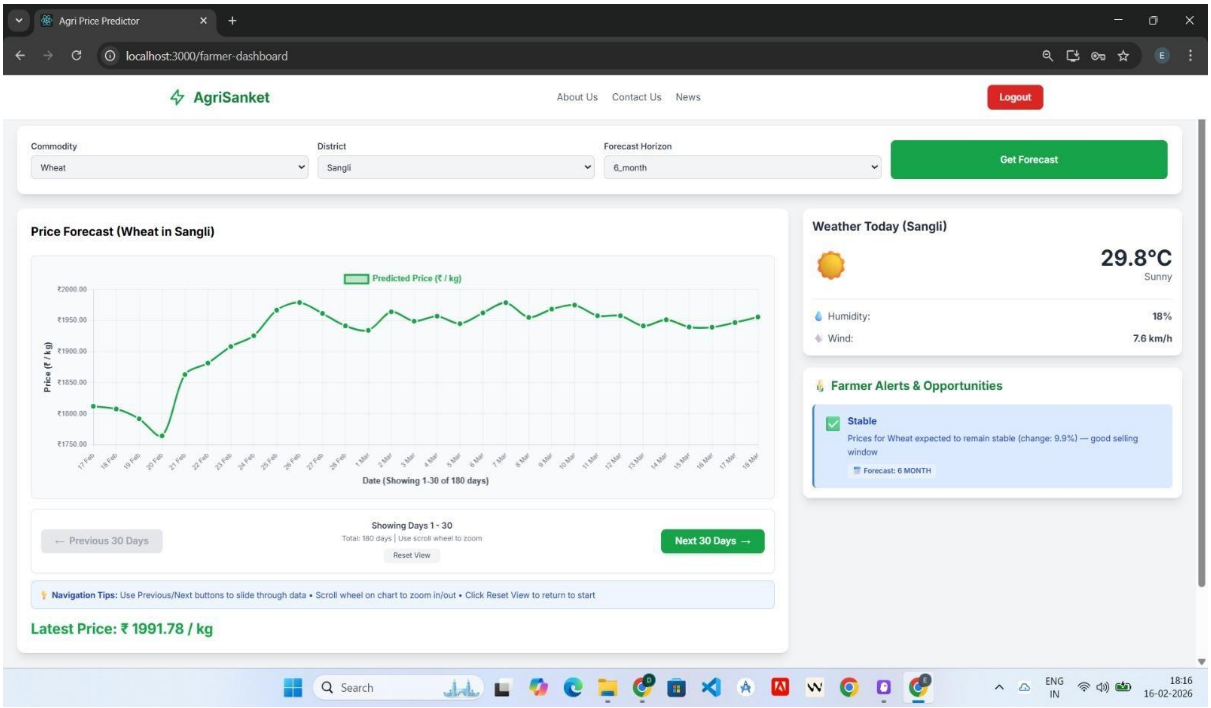

FIGURE 7: Dashboard (Farmer page)

3.4 Ensemble vs. Individual Model Performance

The ensemble approach demonstrates significant improvements across all metrics, with RMSE reduced by approximately 49.6% compared to the baseline ARIMA model. This validates the effectiveness of combining multiple models for agricultural price prediction.

IV. CONCLUSION

This research presents **AgriSanket**, an AI-driven agricultural price prediction platform that successfully addresses commodity price volatility challenges across Maharashtra by integrating historical mandi prices, meteorological data, and temporal features to forecast prices for 22 essential commodities across 10 districts.

Employing a meta-ensemble approach combining ARIMA, XGBoost, and LSTM models, the system achieves superior predictive accuracy exceeding 95% ($R^2 > 0.95$) for 89.5% of commodity-district combinations, with MAPE values ranging from 1.44% to 8.63% depending on commodity volatility. The platform delivers role-specific insights through tailored dashboards: farmers receive personalized selling recommendations and price alerts, policymakers access comprehensive volatility heatmaps for intervention planning, and consumers gain market transparency for informed purchasing decisions.

Experimental validation on 3,623 data points per commodity-district pair confirms the system's robustness and significant improvement over individual baseline models, with the ensemble approach reducing RMSE by 49.6% compared to ARIMA and 17.1% compared to the best-performing individual model (LSTM).

Future enhancements include:

- Satellite-based crop monitoring for real-time yield estimation
- Pan-India expansion to cover all major agricultural states
- Sentiment analysis integration from news and social media
- Mobile application development with push notifications
- Regional language support for rural accessibility

- Incorporation of macroeconomic indicators (inflation, MSP, export policies)

These enhancements position AgriSanket as a comprehensive nationwide agricultural market intelligence solution contributing to enhanced food security, farmer welfare, and economic stability.

ACKNOWLEDGMENT

We would like to express deep gratitude to our guide, **Prof. Reshma Chaudhari**, Department of Computer Engineering, for her constant encouragement and valuable suggestions. The work presented here is possible because of her timely guidance.

We extend our thanks to the panel of examiners—**Prof. Kirtida Naik, Prof. Akshata S. Raut, and Prof. Mitali Ghadigaonkar**—for their time, effort in evaluating our work, and their valuable suggestions.

We thank the Project Coordinator of the Department of Computer Engineering, **Prof. Kirtida Naik**, for her support and coordination, and the Head of the Department of Computer Engineering, **Dr. Ashwini Save**, for her support and coordination.

We are also grateful to the teaching and non-teaching staff of the Department of Computer Engineering who lent their helping hands in providing continuous support.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper

REFERENCES

[1] F. Sun, D. D. (2023). Agricultural price forecasting review paper.
[2] N. Singh, D.D. (2025). Impact of Weather Anomalies on Vegetable Prices in India.
[3] G. K. Van, D.D. (2024). “Forecasting of Retail Price Rise in Tomato, Onion and Potato Commodities in India”, Indian Journal of Agricultural Economics 2023.
[4] G. Avinash, D.D. (2023). “Price Forecasting of TOP (Tomato, Onion and Potato) Commodities using Hidden Markov-based Deep Learning Approach”, Statistics and Applications, vol. 22, no. 2.
[5] H. Cheng, D.D. (2024). Intelligent Crop Price Prediction Model Based on ARIMA and Support Vector Machine.
[6] R. L. Manogna, D.D. (2025). “Enhancing Agricultural Commodity Price Forecasting with Deep Learning”, Scientific Reports.
[7] Z. Chen, D.D. (2022). “Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques”, ASTESJ,2022.
[8] E. E. Okere, D.D. (2025). “Deep Learning-Based Prediction and Forecasting of Tomato Prices for the Cape Town Fresh Produce Market,” Forecasting, 2025.
[9] C. Sun, D.D. (2024). “A Study on Agricultural Commodity Price Prediction Model Based on Secondary Decomposition and Long Short-Term Memory Network,” Agriculture, 2024.
[10] R. L. Manogna, D.D. (2025). “A Novel Hybrid Neural Network-Based Volatility Forecasting of Agricultural Commodity Prices: Empirical Evidence from India,” Journal of Big Data, 2025.
[11] G. H. Harish Nayak, D.D. (2024). “Exogenous Variable Driven Deep Learning Models for Improved Price Forecasting of TOP Crops in India,” Scientific Reports, 2024.
[12] L. Jácome Galarza, D.D. (2025). “AgriTransformer: A Transformer-Based Model with Attention Mechanisms for Enhanced Multimodal Crop Yield Prediction,” Electronics, 2025.
[13] S. Bayona-Ord, D.D. (2021). Machine Learning for Price Prediction for Agricultural Products.
[14] S. Kamble, D.D. (2024). Agricultural Commodities Price Prediction, G. H. Raisoni College of Engineering.
[15] L. Zhang, D.D. (2024). Avocado Price Prediction Using a Hybrid Deep Learning Model: TCN-MLP-Attention Architecture.