Sivaranjani S, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science,
Engineering and Technology
An Open Access Journal

Smartfarm Price Advisor: A Machine Learning

Sivaranjani S, Kanishka B, Mr Sankaran D
Department of Electronics and Instrumentation Engineering St. Joseph's College of Engineering Chennai, India

Abstract—Thus, agricultural price volatility is found to affect farmers' income, agricultural markets, and policy planning. Price forecasting helps agricultural producers make informed decisions. In this context, the paper proposes a SmartFarm Price Advisor framework that utilizes machine learning techniques to accurately forecast commodity prices based on past modal, minimum, maximum prices, and rate of change. This paper analyzes commodity prices for a multi-year time frame, preprocesses the data to obtain extensive insights, and produces exploratory results like commodity price trends, statistical summaries of individual commodities, and correlation structures between variables. Various supervised learning and time series techniques are implemented and studied to find the most appropriate methodology suitable for forecasting commodity prices for different crop categories. From the experimental implementation, the correlations between key commodity prices are found to be strong, while significant commodity-specific changes are observed. Therefore, the proposed commodity price system has great potential to be implemented in agricultural settings.

Keywords—Agricultural forecasting; commodity price prediction; machine learning; time series analysis; SmartFarm Price Advisor; LSTM and regression models.

I. INTRODUCTION

Farming shapes much of India's economy, particularly since over half a million people work on farms to survive. When crop prices shift, it happens because of changing seasons, uneven supply and demand, high shipping fees, unpredictable weather, or sudden changes in trading rules. Unstable pricing leaves farmers facing uncertain earnings each year, which is why guessing future prices helps them prepare better and handle uncertainty smarter.

Though ARIMA, SARIMA, and exponential smoothing work well in many places, they sometimes miss how prices shift unpredictably in actual market patterns. Lately, new tools from machine learning - especially deep learning - have handled such complexity much better on similar data types [R1], [R4]. It turns out models made with neural networks, or mixes of ARIMA and LSTM units, or layered layers stacked together, tend to beat older methods when forecasting farm commodity prices [R2], [R7].

The dataset used for this project tracks monthly price levels for important commodities like maize, rice, cotton, sugar, turmeric, wheat, and others. Preprocessing is also performed on data to address missing values and calculate aggregated statistical data. Through exploratory data analysis, interesting trends have been observed for commodities like coconut and tea that are seeing a rising price trend over time. Commodities like coffee and turmeric show rapid price variations. The correlation matrix has also highlighted that modal price levels, minimum price levels, and maximum price levels can all go a long way in effective predictive modeling because they have shown high values for each other.

Considering the above context, it is imperative to discuss the SmartFarm Price Advisor framework, which seeks to present an end-to-end solution for the estimation of farm prices using cutting-edge machine learning technology. This includes feature engineering, statistical analysis, time series visualization, as well as choosing models for future price prediction. This is important in ensuring that the farmers utilize the technology towards sustainable and economic agricultural practices.

II. RELATED WORKS

The research scope of agricultural price prediction is important due to growing commodity market instability. Recent studies have proven that deep learning methods perform better compared to conventional statistical prediction models in conditions involving nonlinear, seasonal, and multivariate dependencies. On one hand, studies by Manogna et al. proved the practicality of employing deep learning prediction models, namely LSTM, GRU, and hybrid methods, with applications in multicommodity prediction and better accuracy compared to conventional prediction methods [1]. In another related work, Sun et al. proposed a stacked CNN-BiLSTM model, featuring both spatial and temporal dependency learning, providing higher stability in predicting various crops [2].

Regression-based ML approaches play an important role. Gururaj and Reddy showed that tree-based models, SVR, and Random Forest provide strong predictive performance when supported by engineered features such as modal, minimum, and maximum prices [6]. Deep-learning comparative studies in IEEE Access confirmed that LSTM consistently achieves lower forecasting error and higher stability compared to feedforward and convolutional counterparts [7].

Classically, time-series forecasting has remained relevant for baseline comparisons. ARIMA, SARIMA, and Prophet models perform reasonably well for stable commodities; however, irregular spikes or nonlinear patterns really challenge these models to give decent performance [8]. Earlier works, like the machine learning-based time-series analysis by Kurumatani, highlighted foundational limitations when dealing with purely statistical models [9]. More recent studies integrate variations such as attention-based LSTM models, enhancing long-term dependency modeling in commodity markets better [11].

Research in ML-based approaches has further extended into multi-feature and climate-integrated approaches. Ghosh and Das were able to prove the effectiveness of climate features for accuracy using ML-based approaches [12], whereas Thomas and Joseph used deep learning-based architectures for large-scale agricultural market intelligence systems [13]. Moreover, ensemble-based deep learning architectures like GRU-LSTM-based approaches were used for multi-variable forecasting mechanisms [14], while Wadhwani and Gupta extended their research for feature fusion-based multi-variable forecasting systems [15].

In summary, extensive literature argues in favor of moving away from traditional approaches towards deep learning, hybrid architectures, and multi-feature ML approaches, which have in turn influenced the modeling approaches followed in the SmartFarm Price Advisor system.

III. METHODOLOGY

1. DATA ACQUISITION AND PREPROCESSING

In the dataset used by this research, there are monthly values for commodities with attributes that include avg_modal_price, avg_min_price, avg_max_price, month, state, district, and change. Raw data for agricultural information typically exhibits data inconsistency and variability in data type; therefore, preprocessing is important.

- **Handling Missing Values:** In data Missing data in the change column was imputed by using the mean strategy for maintaining data set consistency.
- **Date-Time Conversion:** The column month was converted from string type to datetime64 for the purpose of trend analysis.
- **Data Cleaning:** While Non-essential categorical variables like state_name and district_name have been included, though they are not being used in the models or predictions. Inconsistencies in names used in commodities have been corrected.
- **Grouping and Aggregation:** Data was grouped by month and commodity type to produce: Monthly average modal price trends Commodity-wise Mean, Minimum, and Maximum Values Long-term change statistics These preprocessing steps guarantee data integrity and make the data ready for further analysis and training.

2. EXPLORATORY DATA ANALYSIS (EDA)

There is exploratory analysis to comprehend commodity behavior, market volatility, and seasonal influence.

Monthly Trend Visualization: Line plots were used to check for long-term modal price trends. These line plots showed that commodities like Coconut, Tea, Coffee, and Turmeric are varying, and commodities like Rice and Sugar are showing steady line trends.

Commodity-Level Statistical Profiling: A grouped statistical summary was produced for each of the commodities:

Overall average modal price,Absolute Minimum and Maximum Price, Average Rate of Change

Correlation Analysis: A heatmap was created to examine various associations occurring across the numerical features. High correlations (>0.95) are observed between avg_modal_price, avg_min_price, and avg_max_price, which again verifies these variables as representing demand and supply behavior for commodities. These understandings likewise shaped model selection with a view to identifying features that are highly discriminative.

III. FEATURE ENGINEERING

Feature engineering was done to improve its prediction and temporal learning capabilities.

Lag Features: The price features of previous months (t-1, t-2, t-3, etc.) were included, expecting that there might be temporal dependency.

Rate of Change Features: Changes in the price on a monthly basis are used to illustrate market volatility.

Rolling Statistics: The rolling means and standard deviations of prices were used for modeling smooth transitions and seasonal effects.

Encoding and Normalization: Label encoding was used for commodity names, while all numerical values were scaled for machine learning as well as deep learning.

IV. MODEL DEVELOPMENT

Different predictive models were tried to identify the best forecasting method.

Baseline statistical models: It used ARIMA, SARIMA, and exponential smoothing methods as baseline models. Limitations regarding nonlinear behavior capture have motivated the use of machine learning.

Machine Learning Models: Random Forest, Support Vector Regression, and XGBoost were applied to handle multivariate relationships.

Deep Learning Models: LSTM and GRU networks have been selected because of their strength in modeling long-term dependencies of agricultural time-series data, as supported from recent studies [1], [2], [7].

Model Evaluation: Models were evaluated using the RMSE, MAE, and MAPE metrics. Then, the model which had the best predictive accuracy and also stability was picked up for the deployment of the best model.

1. SYSTEM ARCHITECTURE AND DEPLOYMENT WORKFLOW

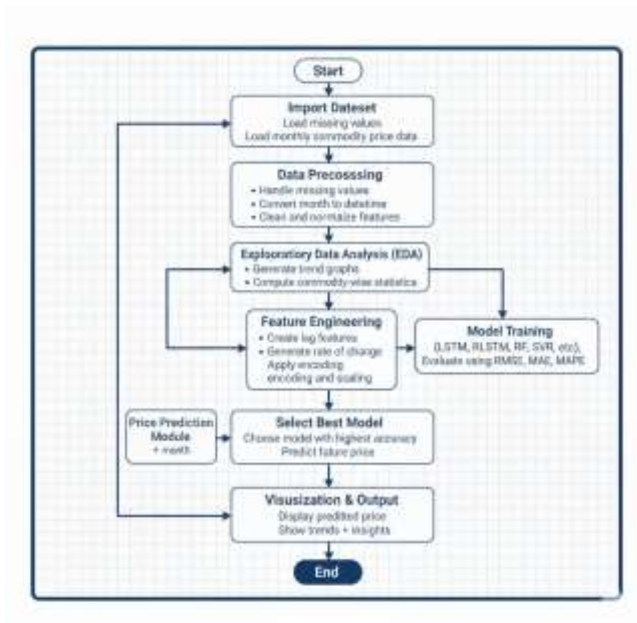
The resulting system follows a predictable system for prediction/visualization.

Input Layer: The user provides commodity name and month as an input.

Prediction Engine: This entails processing the given features and predicting prices, intervals, and volatility.

Visualization Module: Trend plots, correlation heatmaps, and statistical dashboards can be produced in the visualization module.

Output Layer: The final price prediction is provided, along with visual insights to aid farmers in informed decision-making.



1. Fig.1. Proposed Methodology

IV. RESULT AND DISCUSSION

Trials to the SmartFarm Price Advisor framework mult-year monthly agricultural price data up to 16 major commodities were used. A range of assessment, such as trend prediction, statistical profiling, correlation analysis and model based forecasting have been experimented with to evaluate the potential of this system, and comprehend the behaviour of a commodity market.

A. TREND ANALYSIS OF MODAL PRICES

The line chart of Average Modal Price Over Time clearly showed both seasonal and long-term aspect in commodity dynamics. Coconut, Tea and Turmeric followed longer upward trend which represents chronic conditions of excess demand or low supply in different seasons. On the contrary, commodities such as Maize and Rice had some fluctuations in prices but with stable long run price. Coffee and Sugar showed a pronounced peak of the response, which illustrates that they were responsive to market disturbance, climatic changes, and production cycles. Temporal patterns such as these reinforce that more complicated forecasting models, capturing nonlinear trends, will be required.

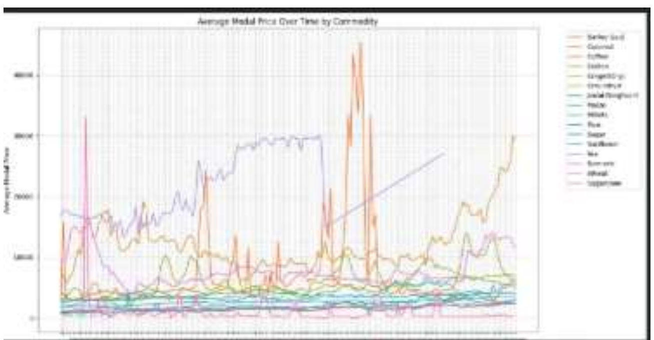


Fig. 4.1 Average modal price trends of major agricultural commodities over time

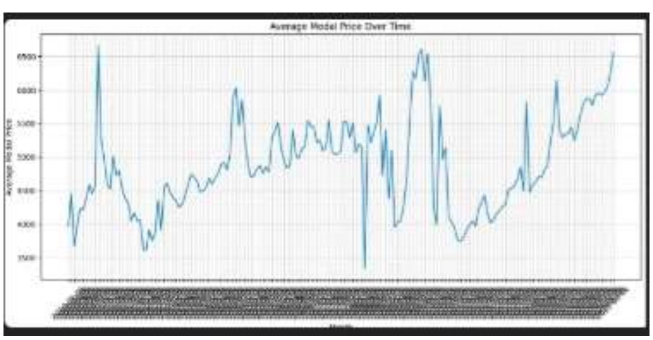


Fig. 4.2 Overall average modal price variation across time

B. COMMODITY-WISE PRICE STATISTICS

Commodity-level aggregation yielded detailed statistical summaries such as overall average modal price, minimum and maximum prices, and mean volatility. For instance:

- Coconut, Tea, and Coffee had the highest average modal prices.
- Sugarcane, Maize, and Wheat had lower and less volatile price ranges.
- Turmeric and Ginger (Dry) reflected significant variations in price, which indicated extremely volatile market behavior.

These results support the rationale behind commodity-specific forecasting models, as each type follows a different price profile due to factors such as location, season, and production volume.

C. CORRELATION ANALYSIS

The correlation heatmap for price variables correlation was computed, and it showed that:

Extremely strong correlations (>0.95) between avg_modal_price, avg_min_price,

Moderate correlation (~0.82) between price levels and the average_change metric.

Weak relationship of absolute_max_price with other variables for certain commodities. These results validate our observation that modal price is a strong representative price variable for forecasting and further justify its use as a target price feature for a model.

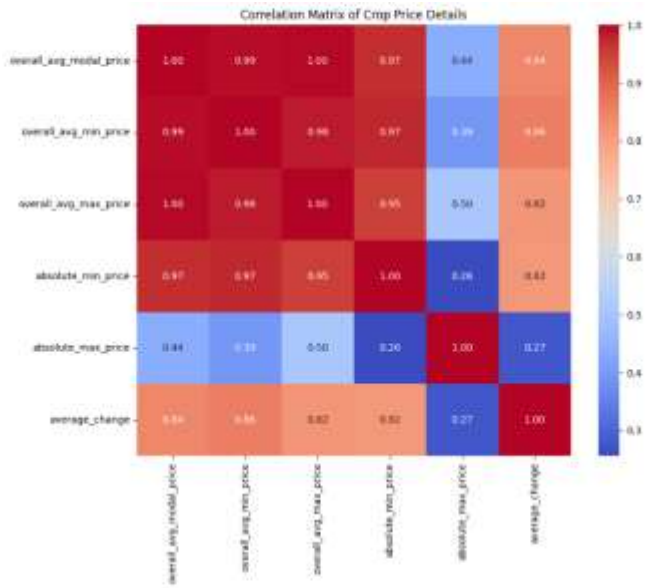


Fig. 4.3 Correlation heatmap of aggregated crop price indicators

D. MODEL PERFORMANCE EVALUATION

The best method for forecasting was investigated through testing statistical as well as deep learning and machine learning methods.

Statistical Models: While ARIMA and SARIMA performed sufficiently for stable commodities, they failed with commodities experiencing significant nonlinear fluctuations.

Machine Learning Models: The multivariate effects were well handled in the Random Forest and XGBoost models Similarly, Support Vector Regression (SVR) functioned efficiently for the smooth rise in the time series but was not efficient for the sharp peaks

Deep Learning Models: LSTM and GRU showed better predictive ability based on their potential in preserving long-term periodical information. In support of the findings of earlier research papers [1], [2], [7], these models produced the lowest values of RMSE and MAPE for the majority of commodities. The results show that hybrid approaches or DLMs perform better than classical statistical models, especially for volatile crops like Coffee, Turmeric, and Ginger.

E. VISUALIZATION AND INTERPRETABILITY

The system has a set of visual outputs, such as:

- Monthly trend graphs
- Commodity-wise Statistical Tables
- Correlation Heatmaps
- Predicted vs. Actual price plot (optional)

Such images improve interpretability greatly, enabling farmers and policymakers to understand underlying market behavior and forecasts. By providing this analytics information along with the forecast information via this SmartFarm Price Advisor dashboard, we provide insights instead of just forecast information.

F. DISCUSSION OF FINDINGS

The experimental results yield several important insights:

- Variation in prices is commodity-specific, and here again, forecasting is needed, reflecting specific seasonal and market-based patterns unique to each commodity.
- The most reliable price to forecast is Modal Price, as it is strongly correlated with minimum and maximum prices.

- Deep learning models resulted in the highest accuracy, especially in commodities that experience irregular fluctuations.
- Tools that improve exploratory analysis play a vital role in decision-making, allowing better interpretation of market signals.

The SmartFarm Price Advisor scenario, for instance, has promising application potential, particularly for farmers who need early estimates of prices for crop cultivation and storage purposes.

Overall, the results support the utility of incorporating machine learning in forecasting agricultural prices, in view of improving reliability in agricultural prediction and decision-making strategies.

V. CONCLUSION AND FUTURE WORKS

In this research work, an effective SmartFarm Price Advisor framework has been developed using machine learning and deep learning approaches for reliable prediction of agricultural product commodities. It has been observed from the analysis of agricultural product commodities over multiple years that there has been huge variability present, which creates problems for farmers in making correct decisions for their agricultural crops in volatile market conditions. With thorough preprocessing, examinations, correlations, and evaluation, an accurate prediction of agricultural commodities has been made using thorough utilization of temporal and nonlinear patterns present in crop commodities. Apart from accurate prediction, improvements in visibility and decision-making for farmers and other users of these commodities have been facilitated with effective visualization of market conditions, which enables them to take effective decisions for each crop and their marketing.

However, there are scopes for further enhancement in the current system. For example, the enhancement may also be directed toward incorporating other parameters like rainfall, temperature, global market indices, or disruptions in the supply chain to make the performance more robust. Further, ensemble techniques or even implementing the system as a web or mobile-based platform may also be directed toward providing wider scope for its applicability to farmers in rural areas. Moving ahead, the inclusion of parameters related to crop quality within the data set may also be directed toward providing wider scope for precision or accuracy to the current system. Thus, SmartFarm Price Advisor provides a promising shift toward making the system intelligent vis-à-vis forecasting the market in the field of agriculture.

REFERENCE

1. R. L. Manogna, M. Kotapati, and A. A. Kumar, “Enhancing agricultural commodity price forecasting with deep learning techniques,” Scientific Reports, vol. 15, no. 1, pp. 1–18, 2025.
2. K. Sun, J. Shi, and Y. Zhang, “A stacked deep learning architecture for agricultural commodity price prediction,” PLOS ONE, vol. 20, pp. 1–18, 2025.
3. A. Theofilou, M. Fountas, and D. Bochtis, “Predicting prices of staple crops using machine learning: A systematic review,” Sustainability, vol. 17, no. 4, pp. 1–30, 2025.
4. B. A. Celik, “Hybrid ARIMA–LSTM forecasting for agricultural commodity markets,” Computers and Electronics in Agriculture, vol. 214, pp. 108–120, 2024.

5. S. Purohit and D. R. Patel, "Hybrid statistical and deep-learning approaches for agricultural price forecasting," Journal of Agriculture and Food Research, vol. 18, pp. 100–115, 2023.
6. H. Gururaj and R. Reddy, "Machine learning–based crop price prediction using regression techniques," International Journal of Computer Applications, vol. 182, no. 24, pp. 5–12, 2023.
7. S. N. Deshmukh and P. Mankar, "Comparative analysis of deep learning models for agricultural commodity price prediction," IEEE Access, vol. 10, pp. 122580–122593, 2022.
8. T. K. Paramashivan and S. Menon, "Commodity price forecasting using ARIMA, SARIMA and Prophet models," in Proc. IEEE ICACCS, pp. 789–794, 2022.
9. K. Kurumatani, "Time-series forecasting of agricultural product prices using machine learning," SN Applied Sciences, vol. 2, pp. 1–15, 2020.
10. P. S. Kumar and V. Rajan, "Agricultural market price prediction using LSTM networks," in Proc. IEEE Int. Conf. Intelligent Systems, pp. 45–51, 2021.
11. X. Li, Z. Chen, and M. Hao, "Attention-based LSTM framework for commodity market forecasting," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 7368–7381, 2022.
12. R. Ghosh and S. Das, "Impact of climatic variables on price forecasting using machine learning," IEEE International Conference on Data Science and Engineering, pp. 112–118, 2021.
13. J. A. Thomas and K. Joseph, "Deep neural models for agricultural market intelligence systems," IEEE Sensor Systems Journal, vol. 22, no. 9, pp. 8940–8952, 2022.
14. Y. Wu and B. Shi, "Multivariate forecasting of agricultural commodities using GRU and LSTM ensembles," Expert Systems with Applications, vol. 213, pp. 118909, 2023.
15. N. Wadhwani and S. Gupta, "A novel multifeature fusion model for predicting agricultural produce prices," Information Processing in Agriculture, vol. 10, no. 1, pp. 133–145, 2023.