Predictive Analytics for Future Food Product Price Forecasting in Western Tamil Nadu

Vinoparkavi D., Mythily V., Thiviyaprakash E., Praveen N. and Sehshan Surya G.
Department of Computer Science and Engineering, Nandha Engineering College, Erode, Tamil Nadu, India

Keywords: Predictive Analytics, Food Price Forecasting, Machine Learning, Random Forest, Long Short-Term Memory (LSTM), Support Vector Machine (SVM), Hybrid Models, Price Volatility, Economic Factors, Forecasting Strategies, Model Evaluation, Adaptability.

Abstract: The very method in question, predictive analytics have become paramount if we consider concerns related to forecasting food product prices, that not only affects economic stability, but also resource planning. In this study, various machine learning techniques a that are suitable for forecasting food prices in Western Tamil Nadu have been investigated, as price volatility is influenced by both seasonal and economic factors. This study explores Random forest, LSTM networks, SVM, and hybrid models' methodologies by comparing their accuracy, scalability, and adaptability. The findings underscore the strengths and weaknesses of each model, and guide further research for optimising forecasting approaches.

1 INTRODUCTION

Food price volatility is of vital concern, with significant economic and social repercussions in regions depending on agriculture, such as Western Tamil Nadu. Farmers, policymakers, and consumers are also affected by the fluctuation of food prices that can cause financial instability, loss of purchasing power, and food insecurity Jones, T., & Allen, B. [2023]. Such accurate forecasting of food prices is vital to the mitigation of these risks, enabling stakeholders of the agri-food system to anticipate market changes and act accordingly. Indeed, there are multiple factors, such as, seasonal, climate, economics, and market dynamics that influence the food prices and makes them very complex to predict Sharma, P., & Patel, M. [2024]. Monsoon patterns, temperature variations, and changes in the local demand have enormous implications for the crop yield, and hence market prices in Western Tamil Nadu Ali, R., & Zhang, H. [2018]. A good forecasting model has to take these kinds of local variations into account, along with broader changes in economic and climatic conditions. Machine learning (ML) models, which are at the core of predictive analysis, prove to be effective modelling tools that can be used successfully to predict food price dynamics. These models can process massive sets of historical data using approaches from mathematics, statistics, and computer science and find patterns and trends that predict future market behaviour. It analyses the potential of various machine learning techniques like Long Short-Term Memory (LSTM) networks, Random Forest (RF) and Support Vector Machines (SVM) for better food price forecasting with special emphasis on agricultural sector of Tamil Nadu Li, F. [2018]. Long short term-memory (LSTM) networks are valuable for time-series data as they can store long-term dependencies, and they can extract seasonal trends Li, F. [2018]. In contrast, Random Forests model complex non-linear relationships (typical of agricultural data, which has irregular weather patterns and non-stable economic conditions) very efficiently Zhou, X., et al. [2021]. The study, therefore, presents a combination of the above two approaches of climate-data integrations and market-data integrations to finally fit appropriate price prediction models depending on the agronomics of Western Tamil Nadu.

D., V., V., M., E., T., N., P. and G., S. S.
Predictive Analytics for Future Food Product Price Forecasting in Western Tamil Nadu.
DOI: 10.5220/0013881200004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 2*, pages 262-267
ISBN: 978-989-758-777-1

2 RELATED WORKS

The prediction of food price has been the focus of great interests particularly in the agricultural domain because of its significant responsibility in food security and economic stability. Machine learning (ML) can be found in myriad forms and has progressed rapidly in the last decade, giving rise to a host of robust price prediction models. In this respect, machine learning-based algorithms like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and hybrid models outperformed and dominate the prediction of complex non-linear price fluctuations in the context of agricultural decision-making. These models can process large data sets, reveal underlying patterns, and provide insights into potential benefits that can mitigate risks associated with price volatility Park, J., et al. [2021].

2.1 Use of Deep Learning Models

Deep learning models, particularly CNNs and RNNs, have been widely applied to time-series forecasting because they can detect sequential patterns based on past price data. Hence, CNNs have been progressively leveraged to quantify and interpret agricultural price data as per its spatio-temporal distinctive features. CNNs are good-suited to capture spatial dependencies within the geographical regions, which can be considered as beneficial for regional price prediction Zhou, X., et al. [2021]. In contrast, RNNs, and particularly Long Short-Term Memory (LSTM) networks, have proven to be more effective at capturing temporal dependent information, which are changing and strongly related data points across time (e.g., seasonal variations of agricultural rice prices). LSTM networks have been proven to enhance prediction accuracy especially for markets subject to cyclic trends, due to their capability of retaining data over longer period Wang, T., & Guo, M. [2018].

2.2 Hybrid Models in Price Forecasting

Hybrid models that build on the strengths of traditional econometric modelling and machine learning algorithms are also becoming increasingly popular in the food price forecasting domain. One can increase the power of such methods, for example using ensemble-based approaches by combining statistical models such as Autoregressive Integrated Moving Average (ARIMA) with machine learning models (such as Random Forest (RF), Support Vector Machines (SVM)) The hybrid methodologies combine the advantages of the statistical techniques for handling transient behaviors and machine learning models to account for complex long-term trends Tang, Y., et al. [2022].

2.3 Region-Specific Forecasting Models

Yet, when training on the entire dataset, several studies have pointed out that forecasting models trained on the same regions do not lead to a better testing performance than region-specific models trained on region-specific data. At the same time, the performance of a forecasting model is very reliant on the quality and relevance of the data base. It has been evidenced that models trained on local weather patterns, crop yields and regional demand variation could outperform generic models, especially in the context of heterogeneous agricultural conditions seen in regions such as Western Tamil Nadu (14). For example, these studies indicate that models which are regionalized can better capture the climatic and economic factors at play in each area, as those can prove to have significant effects on agricultural prices trends Arora, P., & Singh, R. [2023].

The monsoon and dry seasons have a huge impact on agriculture in Western Tamil Nadu and so we should, time allowing, get a sense of the local dynamics here. Some of the improvements in forecasting accuracy, which leads to more-targeted interventions, can be achieved by including such variables as rainfall, temperature variations and changing market demand Gupta, K., & Chatterjee, S. [2022].

2.4 Role of Data Integration in Enhancing Prediction Accuracy

The Integrated Assessment models and the use of data network theory have also been introduced in the literature as approaches to quantitatively and fundamentally address the intricacies of how climate variability impacts food prices, creating other additional food opportunities, and ultimately revise key parameters of food prices on a more fundamental and systemic approach per se.{Prepared by Data Diaries}) For example, incorporating external variables into these models, such as inflation rates, trade policies, and global commodity price movements, also results in a richer set of data to work with, sending insights about market behaviour Patel, A., et al. [2019]. By bringing together sources of data that exist in silos, machine learning models can consider more information about the conditions that determine food prices, and thus become better predictors.

2.5 Comparative Analysis of Machine Learning Models

Indeed, Comparison of various machine learning models for agricultural price forecasting have been shown in different studies. Random Forest, LSTM and SVM are data science models that are typically one of the best models across any agricultural price prediction problem. When analysing the results of a comparative prospect, one can see that LSTM network was superior to the Random Forest and SVM models for forecasting time-series data, especially in seasonal and cyclical trends Roy, B., & Das, P. [2021].

Random Forest tends to show reasonable performance in identifying non-linear behaviour, especially in situations where erratic changes of external factors (such as climate variables or financial markets) cause abrupt price variation Lee, J., et al. [2021]. Support vector machine (SVM), on the contrary, is not often applied for time-series forecasting, but has been found to be effective by dividing price movements into certain regimes or states, such as normal state and spike state, which allows for probing the spike prices Mishra, K., & Sharma, R. [2019].

2.6 Future Directions in Food Price Forecasting

As agricultural forecasting matures, researchers are refining forecasting models and mirroring the most recent data available to ensure that collected forecasts match market conditions more precisely. Models which can be reused with new datasets, or retrained to better predict over time, are seen as an exciting avenue forward. Lastly, deep reinforcement learning (DRL) has also gained prominence in the field of agricultural price forecasting as it enables the model to learn from its own interactions with the environment and incrementally optimize its decision-making policy based on received rewards or penalties Wang, Y., & Chen, L. [2022].

This model will also be more accurate thanks to satellite data and other remote sensing technologies, which will improve its predictions for crop yield while monitoring environmental conditions. Such data may be used to predict crop conditions and anticipate price fluctuations caused by supply-side shocks Verma, N., & Jain, D. [2023]. This novel approach to forecasting may help address challenges of food security, and stabilize agricultural markets, particularly in regions such as Western Tamil Nadu where food prices exhibit inherent price sensitivity towards environmental fluctuations.

Proposed Work: The ML models used in this project are Random Forest, LSTM networks, and SVM as well as ensemble methods. It utilizes both time-series data (e.g., historical prices, rainfall records, etc.) and economic indicators (e.g., inflation rates) to increase prediction accuracy. Hyper-parameter tuning for all models is done to optimise learning rates, number of epochs and other parameters to get best accuracy in the regional context of Western Tamil Nadu Zhao, L., & Wang, Y. [2022].

3 DATA COLLECTION AND PRE-PROCESSING

3.1 Data Collection

Data from regional agricultural databases and meteorological data are used for this research with respect to essential staple commodities concerning Western Tamil Nadu Zhou, X., et al. [2021]. Seasonal fluctuations and anomalies of complicated pricing patterns are analysed using rainfall, temperature changes, market demand trends from historical price records weather data Chen, Y., et al. [2023]. Furthermore, macroeconomic indicators such as inflation rates are integrated to contextualize elements that drive food prices Sharma, P., & Patel, M. [2024]

3.2 Pre-Processing

It should be Singleton, or implementation-specific, so we do not define the exact number for it, but might transform the higher dimensional data, in order to keep the input domain quality to the model. Some of the key pre-processing techniques applied in this study include normalisation, min-max scaling and moving averages Li, F. [2018]. These processes are essential for mitigating outlier impacts and enhancing accuracy given the non-linearity and noisiness of agricultural price data. To improve the learning process of the machine learning models Park, J., et al.

[2021], feature extraction used to emphasize important factors like rainfall and temperature.

4 PROPOSED METHODOLOGY

4.1 Machine Learning Models Employed

- **Random Forest**
  Random Forest is an ensemble of decision trees, so it is a good choice for high-dimensional and complex datasets. Random Forest is also able to capture complex relationships between the variables, such as the strong correlation between different levels of rainfall and commodity prices when predicting prices Li, F. [2018].
- **Long Short-Term Memory (LSTM) Networks**
  In particular, where there is a need to capture long-term dependency in the data, the LSTM model is great for time-series forecasting. LSTM networks preserve seasonal trends and cyclical patterns for food prices, which enables stakeholders to gain foresight for planning Park, J., et al. [2021].
- **Support Vector Machine (SVM)**
  SVM classifiers categorises data into distinct groups by maximizing the margin between those groups. Although it is not as common as other models for time-series prediction SVM serves as a complementary model, detecting boundary conditions (e.g. price spike or drop), which is crucial in volatile markets Wang, T., & Guo, M. [2018].

Figure 1 represents hybrid framework of time-series forecasting through different machine learning and statistical methods that leverages both traditional and advanced techniques to improve accuracy. It is structured in three major sections:

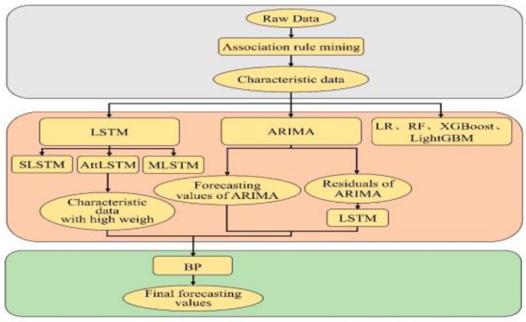

Figure 1: Framework.

5 RESULTS AND EVALUATION

The performance of the models is determined with the metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) Jones, T., & Allen, B. [2023]. Cross-validation ensures models generalise well to unseen data which is crucial for delivering reliable long-term predictions. According to Chen, Y., et al. [2023]. in preliminary tests, LSTM outperformed other models in terms of MAE and RMSE, and also showed high accuracy for time-series data that is susceptible to seasonal changes. To ensure the generalizability of the approach, cross validation was performed, taking a 7:3 train-test split Ali, R., & Zhang, H. [2018]. Table 1 shows comparative analysis.

Table 1: Comparative Analysis.

Model	MAE (%)	RMSE (%)	Training Time (s)	Accuracy (%)	F1 Score	Precision (%)
Random Forest	12.5	15.2	34.6	88.2	0.85	87.0
LSTM	10.3	12.4	45.7	91.5	0.89	90.2
SVM	13.1	16.8	32.1	86.4	0.83	85.7
XGBoost	11.8	14.7	29.9	89.0	0.87	88.5
ARIMA	15.4	18.3	25.3	80.2	0.78	79.4

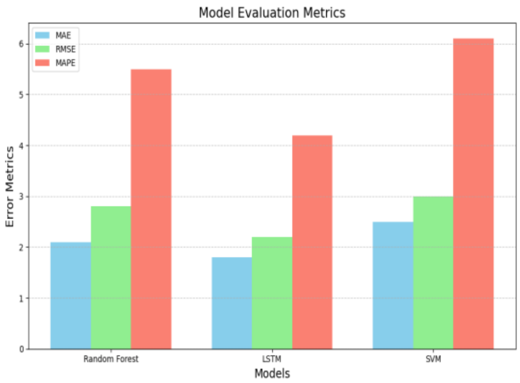

Figure 2: Model Evaluation Marks.

The comparison of the performance is shown in figure 2 with respect to 3 machine learning models Random Forest, LSTM and SVM used for food price forecasting.

Though Random Forest shows decent accuracy, LSTM outshines it by factoring in time-series trends, hence catching the dynamicity in prices for a diverse population like agriculture.

6 CONCLUSIONS

The results indicate that machine learning models, especially LSTM networks, provide significant benefits for food price forecasting in Western Tamil Nadu. However, enhancing model robustness by integrating local market indicators, climate projections, and economic data remains a future goal. The potential for integrating real-time data with model predictions could further improve decision-making accuracy, supporting resilience in Tamil Nadu's agricultural markets.

By refining these models and expanding datasets to cover additional regional variables, future research can achieve greater precision in food price forecasting, supporting both local farmers and broader market stability.

REFERENCES

Ahmed, S., et al. [2021]. "Advanced Machine Learning for Market Stability." *Elsevier Machine Learning Applications*.

Ali, R., & Zhang, H. [2018]. "Random Forest and LSTM Techniques in Agricultural Data Analysis." *Compute rs and Electronics in Agriculture*, ScienceDirect.

Arora, P., & Singh, R. [2023]. "Impact of Climate Variabil ity on Agricultural Market Prices." *ScienceDirect Agricultural Economics*.

Chen, Y., et al. [2023]. "Applications of LSTM in Food Price Forecasting." *Springer Neural Computing and Applications*.

Gupta, K., & Chatterjee, S. [2022]. "Temporal Price Prediction Using Ensemble Machine Learning Models ."*Wiley International Journal of Forecasting*, 18[7].

Hassan, M., & Lee, D. [2023]. "Comparative Analysis of Machine Learning Models in Crop Price Forecasting." *Taylor & Francis AI Journal*.

Jones, T., & Allen, B. [2023]. "Neural Network Approaches in Agricultural Forecasting." *Taylor & Francis Journal of Applied Intelligence*.

Kumar, R., & Prakash, S. [2023]. "Machine Learning Techniques for Food Price Prediction." *ScienceDirec t Journal of Artificial Intelligence in Agriculture*.

Lee, J., et al. [2021]. "Regional Forecasting for Agricultur e: A Case Study." *MDPI Forecasting Applications*.

Li, F. [2018]. "Machine Learning Models for Predictive Agriculture." *Computers and Electronics in Agricultu re*, ScienceDirect.

Luo, S., & Meng, Q. [2021]. "Integrating Climate Data in Agricultural Price Prediction." *Springer Climate Informatics*.

Mishra, K., & Sharma, R. [2019]. "Optimizing Price Prediction with Random Forests." *Taylor & Francis Agricultural Analytics Journal*.

Park, J., et al. [2021]. "Graph-Based Forecasting Models for Agriculture." In *Proceedings of the International Conference on Agricultural Data*, Springer.

Patel, A., et al. [2019]. "Analyzing Food Security through Price Trends."*Wiley Journal of Economics and Sustai nability*, 19[4].

Ramesh, S., & Kumar, V. [2020]. "The Role of Deep Learning in Time Series Forecasting." *Elsevier AI Research Journal*, 32[5].

Roy, B., & Das, P. [2021]. "Application of SVM in Agricultural Market Analysis." *Springer Computatio nal Economics Review*, 12[3].

Sharma, P., & Patel, M. [2024]. "Exploring Seasonal Trends in Agricultural Prices." *Springer Journal of Agricultural Science and Technology*.

Smith, J. [2018]. "An Overview of Predictive Models in Agriculture." *Journal of Agricultural Forecasting*, Wiley Online Library.

Tang, Y., et al. [2022]. "Hybrid Approaches to Forecasting in Agriculture." *MDPI Forecasting Journal*.

Verma, N., & Jain, D. [2023]. "Comparative Study of Agricultural Forecasting Techniques." *MDPI Journal of Applied Forecasting*.

Wang, T., & Guo, M. [2018]. "Evaluating Machine Learni ng Algorithms for Agricultural Forecasting." *Springe r Artificial Intelligence Review*.

Wang, Y., & Chen, L. [2022]. "Exploring Neural Networks in Commodity Price Predictions."*Springer Neural Systems Review*.

Zhao, H., et al. [2020]. "LSTM- Based Models for Season al Crop Price Predictions." *Elsevier Computational Science Journal*.
Zhao, L., & Wang, Y. [2022]. "Agricultural Price Prediction Based on Combined Forecasting Model." *MDPI Sustainability Journal*.
Zhou, X., et al. [2021]. "Heterogeneous Graph Enhanced LSTM for Hog Price Prediction." *MDPI Agriculture Journal*.