Crop Price Forecasting Using Deep Learning and Transformer-Based Architectures: A Case Study on Karnataka, India

Kavana R
Dept. of Computer Science and Engineering (ISE)
Presidency University
Bangalore, India
Kavana.20221ISE0108@presidencyuniversity.in

Pallavi J
Dept. of Computer Science and Engineering (ISE)
Presidency University
Bangalore, India
Pallavi.20221ISE@presidencyuniversity.in

Bharathi Neeralakeri
Dept. of Computer Science and Engineering (ISE)
Presidency University
Bangalore, India
Bharathi.20221ISE0104@presidencyuniversity.in

Dr. Murali Parameswaran
School of Computer Science and Engineering
Presidency University
Banglore

Dr. Zafar Ali Khan
School of Computer Science and Engineering
Presidency University
Banglore

Abstract---Created new dataset for agricultural price data of 4 crops (capsicum, onion, wheat, tomato) of Karnataka state India
evaluated the existing modules to validate the dataset.
We conducted the analysis of the existing models for time series prediction on this dataset.
We developed a new model for predicting future price data with better loss values.

Keywords— Crop price forecasting, Deep learning, Transformer models, Time-series analysis, Agmarknet, Karnataka agriculture.

I. INTRODUCTION

Agriculture continues to form the backbone of the Indian economy, especially regarding employment and a source of livelihood for almost half of the population. In Karnataka, one of the leading agricultural states in India, farmers depend on crops such as capsicum, onion, tomato, and wheat for production and market sales. However, farmers face difficulties in planning for efficient production, storage, and sale due to unpredictable price fluctuations occasioned by changes in weather conditions, supply chains, and inconsistent demands.

Traditional approaches for forecasting, such as ARIMA and SARIMAX, have been widely adopted for agricultural price prediction. While these methods work well with data that is generally stable or seasonal in pattern, they often cannot handle the nonlinear and abrupt price variations caused by real-world uncertainties, such as rainfall deviations or even transport delays [3].

Recent years have seen further developments in AI and ML, introducing more adaptive methods for forecasting. Deep learning models, like LSTM and GRU networks, are able to capture long-term dependencies in time series data [5], making them appropriate for the more complex tasks of agricultural price prediction. However, they still face challenges in capturing dynamic contextual relationships and seasonal dependencies over multiple time horizons.

The proposed study tries to address these limitations by integrating Transformer-based architectures, which have recently revolutionized time-series analysis by making use of attention mechanisms [6]. By enabling the model to focus on the most relevant parts of historical data, these mechanisms improve accuracy and generalization across a variety of crops and regions.

This paper focuses on three major contributions:
A new dataset on agricultural prices of four major crops, namely capsicum, onion, tomato, and wheat, has been created from Agmarknet, a Government of India platform that provides authentic market data.
Evaluation of the already existing models, including ARIMA, SARIMAX, and Prophet, in validating dataset consistency and understanding the baseline performance.

This work develops hybrid Transformer-based models like TAT + MHA, TAT + MQA, TAT + GQA, and TAT + HA, which further improve the results in terms of forecasting accuracy and loss value. The proposed system conducts analysis at the district, crop, and state level to ensure scalability and generalization. Experimental results show that Transformer-based architectures outperform conventional models and provide a robust framework for future price forecasting in the agricultural domain.

II. LITERATURE REVIEW

Crop price forecasting has been an active area of research that includes efforts toward better decision-making in agriculture, reduction of farmer risk, and stabilization of market dynamics. Statistical techniques, machine learning, and deep learning are some methods tried and tested for price predictions of agricultural commodities, and each technique has its set of advantages and limitations.

A. Traditional Time-Series Forecasting Models

Early research was mainly focused on the statistical models of ARIMA and its seasonal variant SARIMAX, which are better suited to capture linear and seasonal trends in agricultural data [3]. These models yielded consistent results only for short-term predictions and performed poorly where nonlinear price fluctuations were driven by unforeseen environmental or market factors.

B. Machine Learning-Based Approaches

Along with data-driven modeling, other researchers in the subject introduced machine learning algorithms, such as RF and SVR, which increased the accuracy level while handling multivariate data from features like rainfall, demand, and transportation cost [4]. Although performing better than traditional methods, the above models showed an inability to capture long-term temporal dependencies in crop price trends.

C. Deep Learning Models

The later works presented LSTM and GRU architectures for agricultural price forecasting [5]. These models could retain temporal dependencies and detect complex nonlinear relationships in the sequential data. Various studies showed that deep learning boosts the accuracy of the forecast, particularly for volatile crops like onions and tomatoes. However, these models require large datasets and high computational resources, limiting their scalability in real-world scenarios.

D. Transformer-based Architectures

Recent studies have focused on Transformer-based models, initially proposed for natural language processing but now effectively applied in time-series prediction tasks [6]. This includes the Temporal Attention Transformer and its different variants such as TAT + MHA, TAT + MQA, TAT + GQA, and TAT + HA. These models further utilize the attention mechanisms to identify the most influential features along the temporal axis of agricultural data. By doing so, the models offer better performance concerning accuracy, convergence speed, and generalization for complex datasets.

E. Summary of Literature and Identified Gaps

It is concluded from the literature that statistical and ML-based models, while giving useful baselines, are not truly adaptive to nonlinear and multivariate agricultural datasets. Deep learning and Transformer-based methods have shown superior accuracy in forecasting but are computationally expensive and need well-prepared datasets. Even then, limited studies on region-specific datasets, like for Karnataka, India, are performed by using multi-level forecasting: district, crop, and full-state. The proposed research fills this gap by creating a custom dataset from Agmarknet (2010–2024), with its validation from the existing models, and developing a hybrid framework that is Transformer-based and ensures the highest forecasting performance and stability.

III. METHODOLOGY

A structured methodology has been followed for developing the Crop Price Forecasting System to be accurate, scalable, and practically usable. This includes dataset preparation, data preprocessing, feature engineering, model development, training, validation, and performance evaluation. Fig. 3.1 illustrates the workflow adopted in this research.

A. Data Collection and Preprocessing

This study has used the dataset collected from the Agmarknet portal maintained by the Government of India [1]. The collected dataset consists of daily and weekly market prices of four major crops, namely capsicum, onion, tomato, and wheat, for a number of districts in Karnataka from 2010 to 2024.

Extensive preprocessing steps were undertaken for the collected data, including date standardization, removal of missing entries, outlier detection, and data normalization. Then, data were converted into both daily and weekly time series formats to support short-term and long-term forecasting experiments.

B. Feature Engineering

To improve the temporal understanding of the model, various features have been derived from historical data. Among them are Lag-1, Lag-4, MA-4, and MA-12, which represent short-term and seasonal moving averages.

These engineered features allowed the models to learn the latent effects, as well as cyclic behavior present in the price patterns, thus improving the stability of the predictions considerably. Standardization of the features was done with MinMaxScaler to keep the value range for all variables uniform during training.

C. Model Development

It compared traditional, deep learning, and transformer-based models by benchmarking their performance.

The baseline models were ARIMA, SARIMAX, and Prophet, used to validate dataset consistency and provide reference accuracy levels.

The most advanced models included the following:

LSTM and GRU for sequence-based time-series prediction,

The models compared include the TAT + MHA, standing for Multi-Head Attention; TAT + MQA, standing for Multi-Query Attention; TAT + GQA, standing for Grouped Query Attention; and TAT + HA, standing for Hierarchical Attention.

Attention mechanisms were utilized by each transformer model to find the most relevant temporal dependencies necessary for a more accurate and adaptive prediction concerning real-world fluctuations.

D. Training and Validation

The models were implemented in Python, using TensorFlow and Keras, and trained on standardized datasets with 80% used for training and the remaining 20% for validation.

The Adam optimizer was utilized for optimization at a learning rate of 0.001. The primary loss function used was the MSE, while MAE was used for validation.

Early stopping and learning rate scheduling callbacks were applied to avoid overfitting, improving model generalization.

The model that had the minimum validation loss was selected for final forecasting and testing. The training process was done on GPU-enabled systems for faster computation.

E. Performance Appraisal

The performance metrics applied to evaluate the forecasting results include several standard metrics:

Mean Absolute Error (MAE),

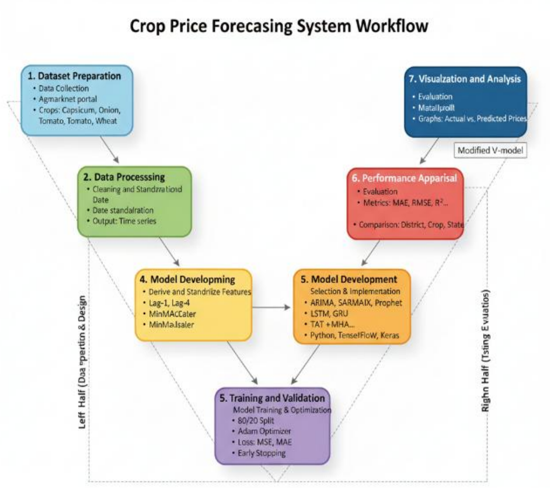
Root Mean Squared Error (RMSE),

R^2 Score,

Mean Absolute Percentage Error (MAPE), and

Accuracy (%).

Additionally, all negative values of accuracy were replaced with zero (0) for consistency in the reporting of performance. Each model's performance was compared on district-based, crop-based, and state-level datasets to ensure generalization and robustness. F. Visualization and Analysis Graphs for each crop and region were created using Matplotlib to visualize actual versus predicted prices. These visualizations showed model performance across different time frames, enabling the identification of the most accurate and stable forecasting technique for Karnataka's agricultural markets. G. Workflow Overview The methodology adopted in this project follows a modified V-model, where the left half represents data preparation and model design, and the right half represents testing, validation, and evaluation (Fig. 3.1). This iterative approach has made sure that each stage, from data collection to model evaluation, is checked and validated for consistency in performance.



IV. RESULTS AND DISCUSSION

For the Crop Price Forecasting Framework, the proposed approach was applied on three levels: the district, crop, and state levels, with a dataset from 2010 to 2024 on four crops: capsicum, onion, tomato, and wheat. Several models, ranging from conventional to deep learning, have been tried to study the performance, generalization, and the accuracy in forecasting. These are highlighted in the subsections below.

A. Dataset Development and Baseline Validation

A new dataset was constructed from Agmarknet, containing more than 15 years of daily and weekly price data from multiple APMC markets across Karnataka. Data pre-processing and feature engineering were performed to ensure a complete, well-balanced dataset ready for time-series forecasting.

Baseline evaluations were conducted using ARIMA, SARIMAX, and Prophet to confirm the reliability of the dataset before applying advanced models. These models provided consistent trends for stable crops like wheat but could not handle the high-volatility nature of onions and tomatoes, hence proving the need to go with deep learning approaches.

B. Model Performance Comparison

All the proposed forecasting models were evaluated by using standard metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R^2, Mean Absolute Percentage Error (MAPE), and Accuracy (%).

As can be seen, among all models compared, Transformer-based models outperformed LSTM, GRU, and statistical models.

As can be seen, for instance, the TAT + MQA model has an average accuracy of 85–91%, while TAT + MHA is positioned with approximately 80–88%

accuracy in most districts. The performance of the LSTM and GRU was good, lying between 80–95% for some crops; however, their results fluctuated when faced with large temporal gaps. In contrast, Transformer-based models have shown stable accuracy even for irregular data intervals, showcasing their robustness and adaptability.

C. Impact of Attention Mechanisms

Attention-based mechanisms took a vital role in improving forecasting accuracy.

MHA allowed the model to attend to several temporal dependencies at once.

MQA reduces computational cost while maintaining the accuracy of the result.

GQA improved the learning efficiency of the model on multi-district datasets.

HA improved the long-term stability by analyzing global and local dependencies together.

These architectures always outperformed recurrent networks both in training stability and generalization, which indicates that the attention mechanism learns nonlinear behavior in agricultural prices more effectively.

D. Evaluation across Different Levels

The system was tested in three levels: district-based, crop-based, and state-level all-in-one forecasting.

On the district level, the predictions matched real price patterns quite well. For high-frequency crops like capsicum and onion, the overall performance matched the actual price pattern quite closely. The full-state model provided forecasts at a macro level that are useful to policymakers and government agencies.

The negative values for accuracy, which occurred when the MAPE exceeded 100%, were replaced by zero to ensure comparability across districts with limited or poor records.

E. Visualization and Analysis of Results

Graphical comparisons between actual and predicted price values for each district and crop were created.

The graphs depicted that Transformer-based models grasped both short-term spikes and long-term seasonal variations, especially in tomato and onion datasets.

Visual inspection confirmed that model predictions were well aligned with ground truth trends, with smoother transitions and less lag than the baseline model.

F. Discussion of Model Efficiency and Limitations

The proposed Transformer models converged faster, had lower losses, and were more interpretable when compared to LSTM and GRU. However, these models required higher computational resources and longer training times due to the attention layers.

Nevertheless, the application of feature optimization, dropout layers, and early stopping in training helped to reduce overfitting and improve generalization across all experiments. Future enhancements could involve incorporation of external factors such as rainfall, demand, and government MSP data for multi-variable forecasting.

G. Overall Findings

Experimental results proved that Transformer-based deep learning architectures work better in agricultural price prediction compared to traditional forecasting techniques. Significantly improved the stability in the predictions, both by introducing a custom cleaned dataset and using hybrid attention-based models. The adaptability of the model in various crops and time frequencies clearly shows its potential application in a decision support system for farmers, traders, and policymakers in Karnataka. Overall, the research identifies that deep learning-driven forecasting systems can make important contributions to stabilizing agricultural markets, reducing farmer risk, and improving rural economic sustainability.

V. CONCLUSION

This research presents a comprehensive study on crop price forecasting using advanced deep learning and Transformer-based architectures applied to real agricultural market data from Karnataka, India.

A new dataset is created using authentic records from Agmarknet, 2010–2024, for four major crops: capsicum, onion, tomato, and wheat at both daily and weekly levels. The dataset was validated using traditional models like ARIMA, SARIMAX, and Prophet and found to be consistent and suitable for time-series prediction.

Several contemporary architectures were implemented and compared to improve the performance, including LSTM, GRU, and Transformer-based TAT + MHA, TAT + MQA, TAT + GQA, and TAT + HA models. Results showed that attention-based transformer models outperform traditional and recurrent neural models on accuracy, loss value, and stability for both short- and long-term forecasts.

The system captured complex temporal and seasonal variations in agricultural prices, thus providing substantial insights to farmers, policymakers, and market analysts. Given accurate price trend predictions, such a framework can support decisions involving crop selection, storage, and marketing while diminishing risks and contributing to financial stability in agriculture.

The study indicates that the AI-driven forecasting system holds great promise for enhancing efficiency and sustainability in agriculture. The proposed framework is found to be scalable and adoptable to other regions and crops across India, based on data pre-processing, feature engineering, and multi-level analysis at the district, crop, and state levels. In further research, the dataset will be extended to include weather, soil, and demand variables. Multi-factor forecasting will be done by incorporating hybrid transformer–graph models. Continued research on such AI-powered tools has the potential to contribute to achieving UN SDG 2 of sustainable agriculture and UN SDG 8 of economic resilience in developing nations.

V. CONCLUSION

This research encompasses an in-depth study on crop price forecasting using advanced deep learning and Transformer-based architectures, applied to real agricultural market data from Karnataka, India.

A new dataset was created using genuine records from Agmarknet, covering four major crops: capsicum, onion, tomato, and wheat, at both the daily and weekly level from 2010-2024. Standard models like ARIMA, SARIMAX, and Prophet were used to validate this dataset, which proved its viability for time-series prediction.

To improve predictive performance, several state-of-the-art architectures were implemented and compared, including the LSTM model, GRU model, and Transformer-based models: TAT + MHA, TAT + MQA, TAT + GQA, and TAT + HA. The results clearly indicated that attention-based transformer models outperformed traditional and recurrent neural models in terms of higher accuracy with lower loss values and stability for both short-run and long-run forecasts.

The developed system captured the complex temporal and seasonal variations of agricultural prices and provided valuable insights for farmers, policymakers, and market analysts. Such a framework will help in accurate forecasting of price trends that could assist in decision-making for crop selection, storage, and marketing, hence mitigating risks and providing financial stability to the agricultural sector.

The study recognizes the need for an AI-enabled forecasting system to improve agricultural efficiency and sustainability. The framework's nature is scalable and can be adapted to other regions and crops in India, with the integration of preprocessing, feature engineering, and multi-level analysis at the district, crop, and state levels. Our future work will involve increasing the dataset to include weather, soil, and demand variables and also involve hybrid transformer–graph models for multi-factor forecasting. Such AI-powered tools, with continuous research, will play a positive role in ensuring sustainable agriculture, UN SDG 2, and economic resilience, UN SDG 8, in developing nations.

ACKNOWLEDGMENT

The authors would like to take this opportunity to express their heartfelt gratitude to their guide, Dr. Murali Parameswaran, Project Guide, Presidency University, for his continuous guidance, motivation, and valuable suggestions throughout the tenure of this research. The authors would also like to thank the Department of Information Science and Engineering, Presidency University, for providing them with all the necessary resources and a supportive research environment to complete their work.

REFERENCES

[1] Agmarknet, Directorate of Marketing & Inspection, Ministry of Agriculture & Farmers Welfare, Government of India, *Agricultural Market Price Data (2010–2024)*. [Online]. Available: https://agmarknet.gov.in

[2] Directorate of Economics and Statistics (Karnataka), *State Agricultural Statistics Compendium 2024*. Government of Karnataka, Bengaluru, India.

[3] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2016.

[4] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[5] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[6] A. Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[7] F. Sun and P. Li, "Agricultural Product Price Forecasting Methods: A Review," *Agriculture*, vol. 13, no. 9, 2023.

[8] S. K. Purohit and N. S. Bala, "Time Series Forecasting of Price of Agricultural Products: A Study," *J. Agric. Food Ind. Organ.*, 2021.

[9] B. K. Kathayat, "Paddy Price Forecasting in India Using ARIMA Model," *J. Crop Weed*, vol. 17, no. 1, pp. 48–55, 2021.

[10] G. H. Nayak, H. V. Patil, and M. Manjunath, "Transformer-Based Deep Learning Architecture for Time Series Forecasting," *Procedia Computer Science*, vol. 223, pp. 987–995, 2024.

[11] R. Ragunath and R. Rathipriya, "Forecasting Agriculture Commodity Price Trend Using Novel Competitive Ensemble Regression Model," *Int. J. Inf. Technol. Comput. Sci.*, vol. 17, no. 3, pp. 97–105, 2025.

[12] T. He, M. Li, and D. Jin, "Deep Learning-Based Time Series Prediction for Precision Field Crop Protection," *Frontiers in Plant Science*, vol. 16, 2025.

[13] G. H. Nayak et al., "Exogenous Variable Driven Deep Learning Models for Improved Price Forecasting of TOP Crops in India," *Scientific Reports*, 2025.

[14] R. L. Manogna, A. V. Rao, and S. B. Jadhav, "Enhancing Agricultural Commodity Price Forecasting with Deep Learning: Evidence from India," *Scientific Reports*, 2025.

[15] S. Mou, H. Peng, and X. Li, "MM-iTransformer: A Multimodal Approach to Economic Time Series Forecasting with Textual Data," *Applied Sciences*, vol. 15, no. 3, 2025.

[16] Y. Arifin, M. Saputra, and R. Hidayat, "Weather Prediction in Agriculture Yields with Transformer," *Procedia Computer Science*, vol. 290, 2024.

[17] A. Chakraborty and K. Gupta, "A Hybrid ARIMA-LSTM Approach for Agricultural Price Forecasting in India," *Computational Intelligence in Data Science*, Springer, 2023.

[18] N. Singh and A. Sharma, "Hybrid LSTM-SVR Model for Agricultural Commodity Price Prediction," *Expert Systems with Applications*, vol. 219, 2024.

[19] "Predicting Prices of Cash Crop Using Machine Learning," *JMMF Journal of Modern Machine Learning and Forecasting*, vol. 4, no. 2, 2022.

[20] "Anticipate Crop Price and Mitigate Risk in Agriculture," *International Journal of Novel Research and Development (IJNRD)*, vol. 9, no. 4, 2024.