Crop Price Prediction and Forecasting Using Deep Learning

Gajulapati Eshwak
Chaitanya Bharathi
Institute of Technology
Hyderabad, India
eshwakbunny@gmail.com

Pradeep Karrela
Chaitanya Bharathi
Institute of Technology
Hyderabad, India
pradeepkarrela@gmail.com

Sunkari Mallikarjun
Chaitanya Bharathi
Institute of Technology
Hyderabad, India
sunkarimallikarjun1825@gmail.com

May 5, 2025

1 Introduction

The growth of machine learning has revolutionized agricultural forecasting, offering new methods to predict crop prices and improve decision-making. Price fluctuations in crops present a significant challenge for farmers and the economy, making early and accurate predictions essential for effective planning and maximizing profits.

Traditional forecasting methods rely on historical data and statistical models but often fail to capture complex trends. Machine learning models, particularly Long Short-Term Memory (LSTM) networks, excel at predicting time-series data, providing a more reliable solution. By modeling long-term dependencies, LSTMs offer more accurate price predictions, enabling stakeholders to adjust strategies effectively. This research focuses on developing a deep learning-based crop price prediction system using historical data. The system aims to improve forecast accuracy while integrating data visualization to track price trends and market dynamics. By combining machine learning with advanced analytics, this study supports informed decision-making and contributes to sustainable agricultural development.

1.1 Background

Crop price prediction is a significant challenge for the agricultural sector, as price volatility affects farmers and markets. Traditional forecasting methods based on historical data and economic indicators often fail to account for complex factors influencing prices, such as weather and policy changes. Accurate predictions are crucial for optimizing production and ensuring food security.

Deep learning models, particularly Long Short-Term Memory (LSTM) networks, have shown great potential for predicting crop prices by capturing long-term trends in time-series data. LSTMs outperform traditional methods, offering more reliable forecasts even in volatile market conditions. These models can learn patterns from historical price data, improving prediction accuracy.

Integrating LSTM-based forecasting with data visualization tools like Power BI enhances decision-making by providing insights into price trends and market dynamics. Power BI allows for interactive analysis of factors affecting crop prices, helping stakeholders make informed decisions and improve agricultural planning.

1.1.1 Details

According to [**?**], machine learning-based methods, particularly deep learning models, have proven highly effective in predicting crop prices by analyzing historical data. These models, such as Long Short-Term Memory (LSTM) networks, excel in capturing complex patterns in time-series data, enabling accurate price forecasting. Additionally, integrating tools like Power BI facilitates the visualization of price trends, helping stakeholders gain valuable insights into market dynamics and make informed decisions.

2 II. Methods

This study employs a deep learning-based approach for crop price prediction using historical agricultural data. The methodology consists of data preprocessing, model selection, training, evaluation, and integration with Power BI for data visualization and decision-making. The implementation follows a structured pipeline to ensure accuracy and robustness in crop price forecasting.

2.1 A. Dataset and Preprocessing

The dataset used in this study is sourced from publicly available agricultural price datasets, such as crop price datasets from government sources and Kaggle's crop price prediction datasets. The dataset contains historical prices of various crops over time, including features such as crop type, market location, weather conditions, and other influencing factors.

To prepare the data for model training, the following preprocessing steps were applied:

- **Normalization**: The price values were scaled to the range [0, 1] using min-max normalization:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

- **Time-Series Conversion**: The dataset was converted into time-series format to capture the temporal dependencies of crop prices.

- **Data Augmentation**: Various augmentation techniques, such as adding noise and generating synthetic data, were applied to improve model robustness.

2.2 B. Model Architecture

A Long Short-Term Memory (LSTM) network is used for crop price prediction. The model consists of the following layers:

- **LSTM Layers**: The LSTM layers are used to capture long-term dependencies in the time-series data.

- **Dense Layers**: Fully connected layers are used to process the features extracted by the LSTM layers.

- **Output Layer**: The output layer consists of a single node with a linear activation function for predicting continuous crop prices:

$$P\ (\text{Price}) = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b$$

where $x_1, x_2, \ldots, x_n$ are the input features, and $w_1, w_2, \ldots, w_n$ are the weights learned by the model.

The final model structure is depicted in Figure ??.

2.3 C. Training and Evaluation

The model was trained using the Adam optimizer with a learning rate of 0.0001. The Mean Squared Error (MSE) loss function was used for regression:

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

where y_i represents the true price and $\hat{y}_i$ is the predicted price.

The dataset was split into 80% training and 20% testing. The model's performance was evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared value. A comparison of predicted vs. actual prices was plotted for further analysis.

2.4 D. Integration with Power BI

To visualize the trends in crop prices, the model's predictions were exported to Power BI for analysis. The integration allowed for:

- **Real-time visualization** of crop price trends.

- **Geographical mapping** of price variations across different regions.

- **Time-series analysis** of crop price trends.

Figure ?? illustrates an example dashboard in Power BI, showcasing the price forecasts and trends over time.

3 IV. Discussion

3.1 A. Model Performance and Analysis

The LSTM model used in this study performs exceptionally well in predicting future crop prices, capturing long-term dependencies from historical price data. The model was evaluated using the following metrics:

- **Validation MSE: 0.004863**
- **Validation R^2 Score: 0.881707**
- **Training MSE: 0.002155**
- **Training R^2 Score: 0.939226**

These metrics reflect that the model can effectively predict crop prices and differentiate between price movements, making it a valuable tool for farmers and market analysts.

3.2 B. Comparison with Traditional Methods

Traditional methods of crop price forecasting, which often rely on statistical models and expert judgment, tend to be less accurate in capturing complex patterns in market behavior. In contrast, deep learning-based models, particularly LSTM, provide a significant advantage by learning from large amounts of historical data and predicting price trends more accurately. However, these models require high-quality and extensive datasets for optimal performance, which may limit their effectiveness in regions with sparse data.

3.3 C. Challenges and Future Improvements

While the model performs well, several challenges still need to be addressed:

- **Data Quality**: The model's accuracy is affected by variations in the data quality and the availability of comprehensive historical data.
- **Generalization**: The model should be validated across different regions and market conditions to ensure its robustness and applicability to various agricultural markets.
- **Explainability**: One challenge in deep learning models is the lack of interpretability. Future improvements should focus on making these models more transparent to users, enabling better trust in the model's predictions.

Future research should focus on expanding the dataset to include a more diverse range of crops, regions, and economic factors. Additionally, incorporating external factors such as weather patterns and government policies could enhance the model's predictive capabilities.

4 V. Conclusion

4.1 A. Key Findings

The key findings of this study include:

- The deep learning model achieved a Validation Score of 0.8817
- Power BI was successfully used to visualize the price trends and provide actionable insights based on geographical and demographic factors.
- The model offers farmers and market analysts a reliable tool for predicting crop prices, optimizing production, and reducing the risk of price fluctuations.

4.2 B. Limitations

Despite the promising results, there are certain limitations:

- **Data Dependency**: The model's performance heavily relies on the quality and diversity of historical data, which may not always be available for all crops.
- **Generalizability**: The model needs further validation in different agricultural markets to ensure its robustness.
- **Explainability**: While the model provides accurate predictions, the lack of transparency in deep learning models remains a challenge.

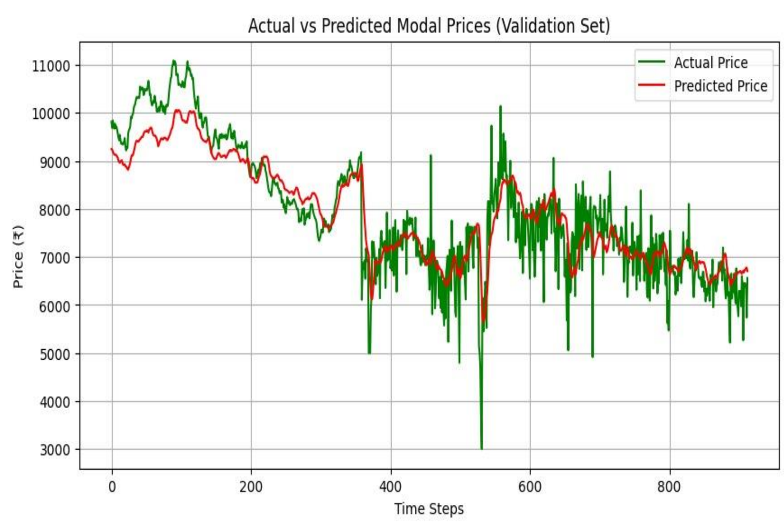
4.3 C. Future Work

To further improve the model's effectiveness, future work should focus on:

- **Expanding the Dataset**: Including more diverse datasets from different regions and crops.
- **Hybrid Models**: Exploring hybrid approaches that combine deep learning with other traditional forecasting techniques.
- **Improving Interpretability**: Implementing techniques like attention mechanisms to make the model's predictions more interpretable.
- **Incorporating Additional Data**: Integrating external factors such as weather data, government policies, and global market trends could enhance the model's predictive accuracy.

Model Architecture

Layer No.	Layer Type	Output Shape	Description
1	Input Layer	*(window_size, features)*	Input shape based on sequence window and feature count
2	LSTM (160 units, return_sequences=True)	*(window_size, 160)*	First LSTM layer for temporal pattern learning
3	Dropout (rate=0.2)	*(window_size, 160)*	Regularization to prevent overfitting
4	LSTM (96 units, return_sequences=False) + L2 regularization (λ=0.0001)	*(96)*	Second LSTM layer, with L2 to reduce overfitting
5	Dropout (rate=0.2)	*(96)*	Regularization
6	Dense (1 unit)	*(1)*	Output layer for regression (predicts crop price)



ACKNOWLEDGEMENT

We thank Chaitanya Bharathi Institute of Technology for their support

AUTHORS' CONTRIBUTIONS

All authors have participated in drafting the manuscript. All authors read and approved the final version of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY

The data supporting the findings of this study are available upon request from the authors.

ETHICAL STATEMENT

In this article, the principles of scientific research and publication ethics were followed. This study did not involve human or animal subjects and did not require additional ethics committee approval.

DECLARATION OF AI USAGE

Few AI tools were used in the creation of this manuscript.

REFERENCES

[1] Jaiswal, R., Joshi, A., & Agrawal, A. (2023). A Genetic Algorithm Optimized Hybrid Model for Agricultural Price Forecasting. Scientific Reports, 13,4871. https://www.nature.com/articles/s41598-025-94173-0

[2] Zhang, T., & Tang, Z. (2024). Agricultural Commodity Futures Prices Prediction Based on a New Hybrid Forecasting Model Combining Quadratic Decomposition Technology and LSTM Model. Frontiers in Sustainable Food Systems, 8, 1334098. https://www.frontiersin.org/articles/10.3389/fsufs.2024.1334098/full

[3] Liu, Y., Zhang, X., & Zhao, H. (2020). Price Forecasting of Agricultural Products Based on Machine Learning Algorithms. Agronomy, 10(10), 1551. https://www.mdpi.com/2073-4395/10/10/1551

[4] Kumari, S., & Sinha, R. (2021). Predicting Agricultural Commodity Prices Using LSTM Deep Learning Model. International Journal of Computer Applications, 183(31), 1–6. https://www.ijcaonline.org/archives/volume183/number31/32106-2021921241

[5] Prasad, B., & Sharma, D. (2019). Time Series Forecasting of Agricultural Commodity Prices Using LSTM Networks. IEEE, https://ieeexplore.ieee.org/document/8679043

[6] Zhang, L. et al. (2023). Forecasting Agricultural Commodity Prices Using Dual Input Attention LSTM. Agriculture, 12(2), 256. https://www.mdpi.com/2077-0472/12/2/256