

Efficient commodity price forecasting using long short-term memory model

Mohammad Tami, Amani Yousef Owda

Department of Natural, Engineering and Technology Sciences, Arab American University, Ramallah, Palestine

Article Info

Article history:

Received Aug 2, 2023

Revised Oct 16, 2023

Accepted Oct 29, 2023

Keywords:

Deep learning

Feature engineering

Long short-term memory

Price instability

Time-series forecasting

ABSTRACT

Predicting commodity prices, particularly food prices, is a significant concern for various stakeholders, especially in regions that are highly sensitive to commodity price volatility. Historically, many machine learning models like autoregressive integrated moving average (ARIMA) and support vector machine (SVM) have been suggested to overcome the forecasting task. These models struggle to capture the multifaceted and dynamic factors influencing these prices. Recently, deep learning approaches have demonstrated considerable promise in handling complex forecasting tasks. This paper presents a novel long short-term memory (LSTM) network-based model for commodity price forecasting. The model uses five essential commodities namely bread, meat, milk, oil, and petrol. The proposed model focuses on advanced feature engineering which involves moving averages, price volatility, and past prices. The results reveal that our model outperforms traditional methods as it achieves 0.14, 3.04%, and 98.2% for root mean square error (RMSE), mean absolute percentage error (MAPE), and R-squared (R^2), respectively. In addition to the simplicity of the model, which consists of an LSTM single-cell architecture that reduced the training time to a few minutes instead of hours. This paper contributes to the economic literature on price prediction using advanced deep learning techniques as well as provides practical implications for managing commodity price instability globally.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Amani Yousef Owda

Department of Natural, Engineering and Technology Sciences, Arab American University

Ramallah P600, Palestine

Email: amani.Owda@aaup.edu

1. INTRODUCTION

Commodity price forecasting is an essential task for stakeholders such as governments, policymakers, retailers, and customers. Accurate commodity price forecasting has a huge impact on managing inflation, securing commodity supplies, and avoiding socio-economy disruptions [1], [2]. Particularly in a world marked by interconnected economies and complex trade relationships [3], precise commodity price forecasts can aid in avoiding crises and promoting stability in the global commodity supply chain [4]. In recent years, commodity price fluctuation has been a significant concern for stakeholders around the world. Several factors have contributed to these fluctuations, leading to challenges in managing commodity supplies, inflation, and socio-economic stability [5]-[8]. Particularly, food price fluctuations can have far-reaching consequences, especially for vulnerable populations in low-income countries, where a significant portion of income is spent on food. High and unpredictable food prices can lead to food insecurity, malnutrition, and social unrest [9], [10]. Governments, policymakers, and international

organizations continue to work on measures to mitigate the impact of these fluctuations. Accurate food price forecasting remains a crucial tool in managing these challenges and promoting global food security [11]-[13].

Many traditional statistics and machine learning (ML) models have been utilized to accomplish commodity price prediction [14]-[18], specifically food price forecasting [19]-[22]. These models are easy to implement and don't require huge computing power. However, these models failed to capture complex patterns and trends in the time-series data since they have prior assumptions about the data [23], [24]. With the rise of deep learning (DL) power, neural network (NN) models gain attraction for being used in many downstream forecasting tasks [25], [26]. Historically, different methods have been applied in commodity price prediction. These include statistical approaches like autoregressive integrated moving average (ARIMA) [27] and seasonal autoregressive integrated moving average (SARIMA) [28], machine learning techniques such as support vector machine (SVM) [29], and deep learning models like long short-term memory (LSTM) [30], [31]. LSTM models are often preferred over other models for time series forecasting due to their ability to capture long-term dependencies and handle sequential data effectively [32]. LSTM's non-parametric characteristic, combined with its ability to handle non-linear patterns and its independence from the need for a stationary process, make it a favorable option for implementation in time-series applications [33].

Various researchers have extended LSTM models to improve forecasting performance. For instance, Ly *et al.* [34] introduced a hybrid model that combines the strengths of LSTM and ARIMA to forecast the prices of cotton and oil, their approach depends on training two separated models, namely LSTM and ARIMA, then averaging the forecasting result from both trained models to achieve better result compared to each individual model. In contrast, Krishnan *et al.* [35] utilized a diverse set of complex LSTM models, namely basic LSTM, bidirectional LSTM (Bi-LSTM), stacked LSTM, convolutional neural network LSTM (CNN-LSTM), and convolutional LSTM (Conv-LSTM), in their five-commodity forecasting study. Each of these models offered unique architectural variations to explore and analyze their predictive capabilities. The authors argued that their complex models are capable of capturing complex patterns and dependencies within the commodity market data, allowing for more accurate and robust predictions. On the other hand, the use of deep learning models in combination with natural language processing such as transformer-based models [36] and bidirectional encoder representations from transformers (BERT) [37] has been gaining traction in recent years due to their ability for capturing market sentiment tracking and context-aware analysis. For instance, Sonkiya *et al.* [38] proposed a generative adversarial network (GAN) combined with BERT model for predicting the price of stock predictions, in their study, BERT is utilized to analyze news and headlines, extracting valuable insights. These insights are then incorporated into the GAN model as external factors, which lead to enhancing their model's predictive capabilities.

Despite the advancements in machine learning and deep learning models, there remains a need for models that are not only accurate but also interpretable and computationally efficient. The black-box nature of many deep learning models often leads to a need for more interpretability, which can slow down their adoption in certain sectors [39]. On the other hand, the computational requirements of these models can be a limiting factor, especially in low-resource settings [40]. This calls for the development of models that balance accuracy, interpretability, and efficiency, which is the focus of this paper. This paper aims to address these limitations by introducing a straightforward effective LSTM model for predicting the prices of various commodities. The developed model demonstrates comparable performance compared to other deep learning models. The paper approach places significant emphasis on the feature engineering aspect of the task. It was found that implementing diverse feature transformations, such as moving averages and volatility changes, greatly enhances the performance of the suggested model. The model's simplicity and efficacy, combined with less demand for computational resources, suggests its potential for broader application in commodity price forecasting. Future work could involve the inclusion of external factors such as social media sentiment and news headlines into the model to further improve its predictive performance.

2. METHODOLOGY

A comprehensive methodology is used in this paper to address the research objectives effectively. Figure 1 illustrates the chronological flow chart of the adapted methodology. The first step involved data acquisition. Subsequently, a data cleaning process was performed where some missing prices were imputed by the average of the lag and lead prices. Next, an exploratory data analysis (EDA) phase was conducted to gain deeper insights into the dataset, identify patterns, and establish descriptive statistics that aid in understanding the underlying characteristics of the data. Following the EDA, feature engineering techniques were employed to transform the raw data into meaningful and informative features that can enhance the performance of the models. After feature engineering, data transformation techniques were applied to normalize the data, ensuring that all variables are on a comparable scale. This step helps in improving the model's convergence and performance. Subsequently, the modeling phase involved a training loop, wherein

various hyperparameters were tested to optimize the model's performance and achieve the best possible results. Finally, the performance of the trained models was evaluated using appropriate evaluation metrics to assess their effectiveness in solving the research problem. This methodology ensures a systematic and rigorous approach to conducting the study, enabling reliable and insightful conclusions to be drawn from the analysis. Through this research, we will illustrate the process of choosing appropriate features, transforming them suitably, and training an LSTM model for forecasting commodity prices. The outcome is expected to provide valuable insights for stakeholders in the commodity industry, enhancing their decision-making capabilities and promoting more efficient practices.



2.1. Dataset

The foundation of the proposed model relies on an extensive dataset that encloses a detailed history of various commodity prices within the State of Palestine. The origin of this dataset is the World Food Program Price Database [41], an extensive global source of commodity price information that spans 98 countries and approximately 3,000 distinct markets. The database audits prices for many commodities, including but not limited to bread, meat, milk, oil, and petrol. Although the database is refreshed weekly, the availability of monthly data is more frequent. Specifically, the dataset under consideration for current research referred to as 'Palestine Food Price Dataset', boasts a rich history of commodity prices in Palestine dating back to 2007, and spans approximately 28,000 entries. The volume of data provided in this dataset introduces an opportunity to derive significant insights into food security studies within this region. The Palestine food price dataset consists of 14 attributes, which include date (on a monthly frequency), district (West Bank & Gaza), city, geographical locations (latitude and longitude), commodity category, commodity item, unit, and price, among others. Samples of prices were collected from twelve distinct cities in the West Bank and Gaza. The dataset comprises commodities that are categorized into eight groups, with a total of 39 categories within the commodity type. The dataset was cleaned and preprocessed prior to analysis. This process involved dealing with missing data, transforming attribute types, and normalizing certain variables.

2.2. Feature engineering

Time series datasets distinguish themselves from other types of datasets mainly due to their inclusion of a temporal component (time), which introduces an extra dimension to the analysis [42]. Unlike typical datasets where data points might be independent, time series data is sequential, and each point often has a relationship with its preceding and sometimes succeeding points. The temporal nature of such datasets makes them particularly challenging to analyze and model, especially when the goal is to forecast future values. Additionally, due to the inherent characteristics of time series data like autocorrelation, the risk of misleading interpretations increases if the analysis isn't handled properly. To address these challenges and effectively capture underlying patterns, it's essential to apply certain transformations. These transformations not only assist in identifying trends and seasonality but also in reducing the impact of noise and outliers. By enhancing each data point with rich information from prior periods, models can be trained to make more accurate and informed predictions, ultimately leading to more insightful and actionable results [43].

2.2.1. Moving average

A moving average, also known as a rolling or running average, is a technique often used in time-series data analysis to smooth out short-term fluctuations and highlight long-term trends or cycles [44]. The idea is to calculate the average of a particular subset of numbers, and as new data comes in, recalculate that average by moving the subset window forward as described in (1).

$$\text{Simple moving average (SMA)} = \frac{\sum_{i=1}^N P_i}{N} \quad (1)$$

Where P_i is the price of the i^{th} period, N is the total number of time periods.

On the other hand, the exponential moving average (EMA) is a type of moving average that places more weight and importance on the most recent data points while still considering the historical data [45]. This makes the EMA more responsive to recent price changes compared to the simple moving average (SMA), which gives equal weight to all data points as shown in (2).

$$\text{Exponential moving average, (EMA)} = (C - P) \times \frac{2}{N+1} + P \quad (2)$$

Where: C is the current data point price, P is the exponential moving average of the previous period, and N is the total number of time periods. * Simple moving average applied for the first period.

In this study, we found that the simple moving average worked better when we used it on four different time periods: 6 months, 12 months, 18 months, and 24 months as illustrated in Figure 2. The superior performance of the simple moving average over the four distinct time periods can be attributed to several reasons. Firstly, commodity prices often follow cyclical patterns and trends, and the simple moving average, being a trend-following method, can effectively smooth out short-term fluctuations, capturing the underlying price movement. Secondly, using multiple time periods allows capturing varying dynamics of the market, from short-term to more extended cycles, providing a more comprehensive insight.

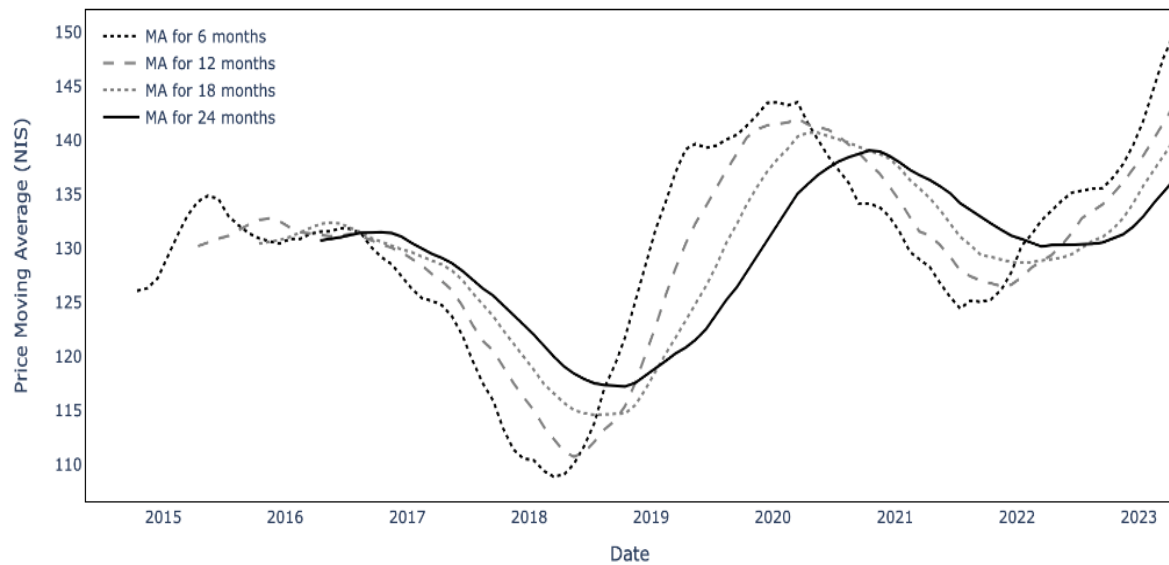


Figure 2. Simple moving average of five window sizes

Results in Figure 2 display four simple moving averages (SMAs) corresponding to window sizes of 6, 12, 18 and 24 months duration respectively. Each SMA line shows the average price of the asset over its respective time frame, providing insights into short, medium, and long-term trends. The 6-month SMA reacts quickly to recent price changes, while the 24-month SMA offers a smoother, long-term perspective, identifying overall trends and potential turning points. Incorporating these different SMAs as features can enhance the proposed LSTM model's ability to capture various time horizons in the data. It allows the LSTM to consider both short-term patterns and long-term trends, improving the accuracy and robustness of the predictions.

2.2.2. Volatility

Price volatility refers to the rate at which the price of an asset, such as a stock or commodity, increases or decreases. It is a statistical measure of the range of the change for a given market index. High volatility indicates that the commodity's price can change significantly in a short time frame in any direction, whereas low volatility implies that the price remains steady [46].

Volatility is often calculated using the standard deviation or variance between returns from the same market index. The most used method is the standard deviation which is typically calculated as illustrated in (3), where n is the number of price returns used in the calculation, x represents each individual price return, μ is the average (mean) of the price returns, Σ denotes the sum of the squared differences.

$$\text{Volatility} = \sqrt{\frac{\sum_{i=1}^n (x - \mu)^2}{n}} \quad (3)$$

The price volatilities for each data point were calculated over five different time frames, including three months (quarterly), six months (semi-annually), one year, two years, and five years. This analysis aids the model in predicting future trends. When there is significant volatility during a certain period, the model is more likely to forecast a greater probability of significant price fluctuations as illustrated in Figure 3.

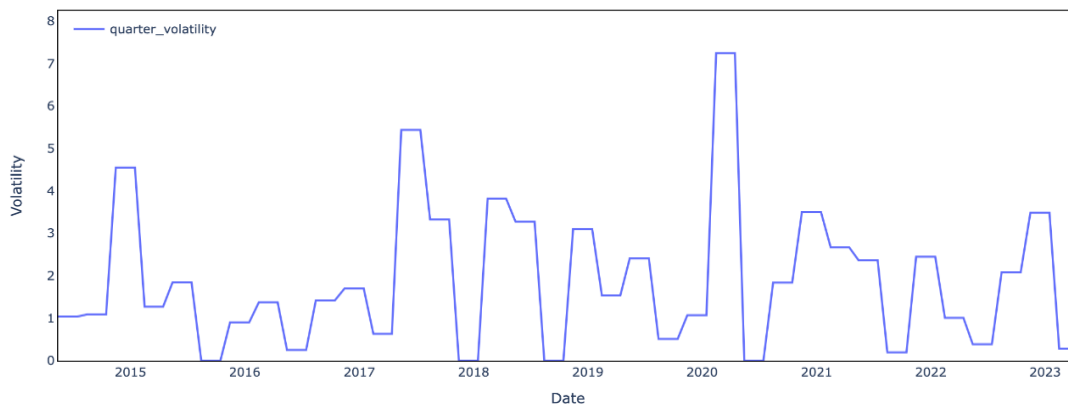


Figure 3. Price volatility over a 3-month window size

Results in Figure 3 show the volatility of a price over a 3-month window size. It shows a line graph with the x-axis representing time and the y-axis representing the calculated volatility of the asset's price at each data point. Volatility measures the level of price fluctuation; higher volatility indicates greater price variability. The chart's line illustrates how the volatility changes over time, highlighting periods of higher and lower price instability. In the task of price forecasting using LSTM models, these volatility lines play a crucial role. LSTM models can incorporate volatility as a feature, helping to capture market dynamics and improve forecasting accuracy. High volatility periods can indicate potential market disruptions or significant price movements, which LSTM models can leverage to generate more accurate predictions.

2.3. Long short-term memory (LSTM): An overview

Long short-term memory (LSTM) is a variation of recurrent neural network (RNN) architecture, designed to model temporal sequences and their long-range dependencies more accurately than vanilla RNNs. It was proposed by Hochreiter and Schmidhuber [32] in 1997. The key to LSTM is the cell state. This is a kind of "conveyor belt" that carries information across time steps with only minimal changes, which helps to mitigate the vanishing gradient problem faced by traditional RNNs [47]. In LSTMs, the information flows through a mechanism that is controlled by various gates as illustrated in Figure 4. These gates decide what information should be kept or discarded at each time step.

- Forget gate: This gate decides which piece of info should be kept or thrown away. Input is passed through this gate, which processes it using a sigmoid function.
- Input gate: The input gate produces a new cell's state via utilizing a sigmoid function that decides which part to be updated and a tanh which creates a new candidate vector.
- Output gate: This gate produces the next hidden state to the cell that embedded info about the previous input. This hidden state is crucial for prediction.

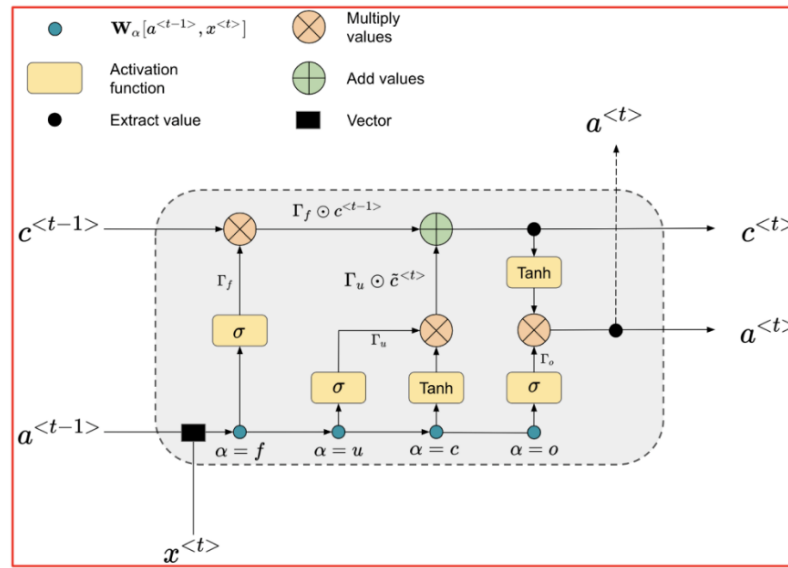


Figure 4. Single-layer LSTM network used for sequence modeling [34]

2.4. Model architecture and training loop

The primary aim of this study is to utilize sophisticated feature engineering methods to build a straightforward yet efficient deep learning model. The LSTM model architecture constructed 77 input layers, and 4 hidden layers, succeeded by a fully connected neural network to combine the spatial data from the surrounding station's layer, which takes 4 inputs from the final hidden layer. Lastly, the rectified linear unit (ReLU) function is applied to prevent the model from predicting negative values. This structure excels due to the advanced transformations applied to the data prior to the training. The training was carried out on a local MAC laptop equipped with 16GB memory and an 8-Core Intel processor. Remarkably, the training time for the model on the specified dataset was a few minutes, which is a significant time reduction compared to the computational resources of other researchers which usually take hours as other researchers reported in [33]-[34]. The LSTM model was implemented using the pytorch deep learning library. Three categorical variables (city, category, commodity) were encoded using one hot encoding. The numerical variables were normalized using MinMaxScaler to ensure they were in the same range. Lastly, the dataset was divided into three subsets: training, evaluation, and testing, with ratios of 70%, 15%, and 15% respectively as illustrated in Figure 5.

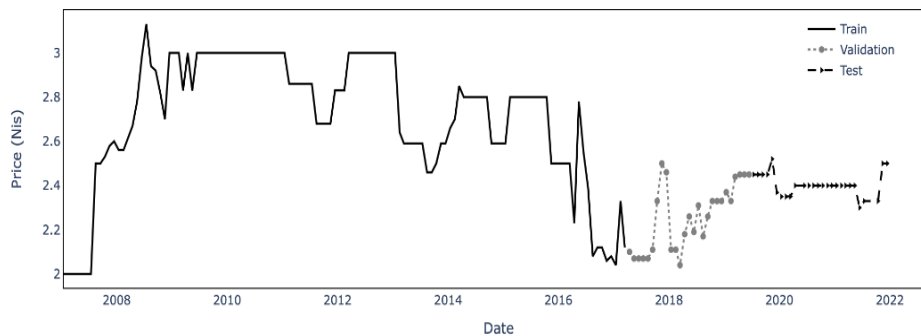


Figure 5. The dataset splits into three subsets: training, evaluation, and testing

Model hyperparameters play a vital role in determining the quality of results obtained from the training process. In this particular scenario, the model was trained using a learning rate of 1e-3, striking a balance between precision and efficiency. To measure the model convergence, the mean squared error (MSE) was employed as the loss function as shown in (4).

$$\text{Mean square error (MSE)} = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2 \quad (4)$$

Where n is the number of data points, Y_i is the observed price, and F_i is the forecasted price.

To optimize the training process, the widely used Adam optimizer [48] was utilized, leveraging its adaptive learning rate capabilities. The training loop was executed for a substantial number of iterations, precisely 1000 epochs, ensuring that the model had the opportunity to learn from the data. To prevent overfitting, two effective early-stopping techniques were employed [49]. These techniques acted as gatekeepers during the training process, monitoring the model's performance closely. Training would halt if either there was minimal improvement in the training loss; precisely 1/1000 of the learning rate or if the validation loss increased for two successive epochs. By incorporating these strategies, as described in Table 1, the training process became more robust and resilient to overfitting, ultimately leading to a more accurate model.

Table 1. Model hyperparameters

Parameter	Value
Epoch	1000
Learning rate	1E-3
Loss function	MSE
Optimizer	Adam
Regularization	Early stop

2.5. Model evaluation

The performance of the proposed model was assessed using three evaluation metrics described in (5) to (7) and those are: root mean squared error (RMSE), Mean absolute percentage error (MAPE), and the coefficient of determination (R^2). RMSE measures the square root of the average of the squared differences between predicted and actual values, providing a measure of the model's accuracy. MAPE calculates the average percentage difference between predicted and actual values, offering insights into the model's relative performance. Lastly, the coefficient of determination (R^2) indicates the proportion of the variance in the dependent variable that is predictable from the independent variable, indicating the model's ability to explain the data's variability. In equation (5)-(7), A_i and F_i are the actual and forecasted prices, respectively, μ_y is the mean price of all data points, and n is the number of data points.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (A_i - F_i)^2}{n}} \quad (5)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left(\frac{|A_i - F_i|}{A_i} \right) \times 100\% \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (A_i - F_i)^2}{\sum_{i=1}^n (A_i - \mu_y)^2} \quad (7)$$

3. RESULTS AND DISCUSSION

The utilized dataset consists of eight categories containing thirty-nine distinct commodities. For training the model, five commodities were selected, resulting in varied but closely related outcomes. The differences in these outcomes can be attributed to the availability of pricing data for each commodity. Among the predictions, the estimation of Bread price showed exceptional performance as shown in Figure 6, with a RMSE of merely 0.14, this suggests that the average difference between the forecasted and actual prices was exceptionally low, indicating the model's precision in capturing the price fluctuations. Additionally, the MAPE of 3.04% indicates that, on average, the model's predictions were within a very small percentage of the actual prices, further substantiating its reliability. Moreover, the high R-squared value of 98.2% denotes an impressive fit of the model to the observed data, indicating that a significant proportion of the variability in Bread prices was accurately accounted by the LSTM model. For other commodities i.e. (meat, milk, oil, and petrol). Table 2 summarizes the evaluations conducted on all of them.

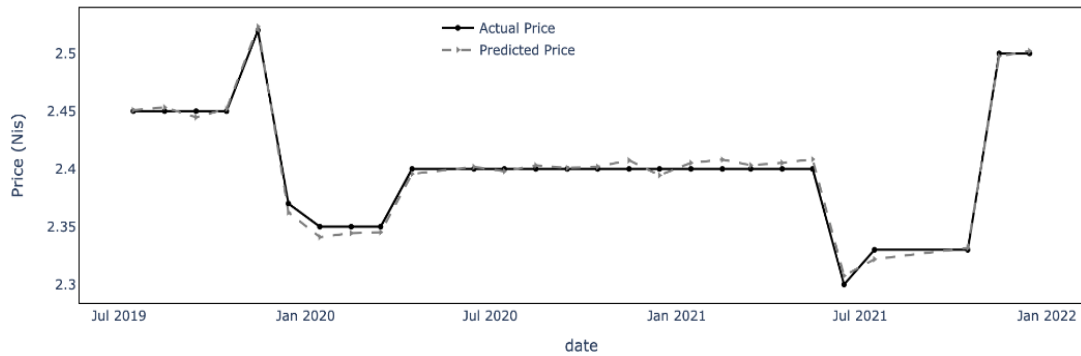


Figure 6. The actual and predicted price of Bread

Table 2. Model evaluation on five commodities

Commodity	RMSE	MAPE	R ²
Bread	0.14	3.04%	98.2%
Meat	0.85	4.15%	80%
Milk	1.2	1.09%	93%
Oil (olive)	0.87	2.26%	90%
Petrol	0.22	2.18%	87.4%

The findings in this paper show a notable advancement of the proposed model over earlier published models. Several previously published models have been compared to our model to offer a thorough assessment of the efficiency and accuracy of the proposed model in this paper. The advantages of the proposed model become clear when considering its superior scores across all evaluation metrics. Furthermore, while many models in the past have shown strength in one or two metrics but weakness in others, our model maintains consistency in its high performance across RMSE, MAPE, and R². This consistency is indicative of the robustness of our approach, setting a new standard in the field. Table 3 presents a comparison of different predictive models in the literature, evaluating their performance based on RMSE, MAPE, and R² metrics.

Table 3. Comparison between results in the literature and the proposed model

Model	RMSE*	MAPE*	R ² *	Reference
Hybrid (ARIMA + LSTM)	0.15	4.3%	-	[34]
Stacked LSTM	0.15	0.079%	96.8%	[35]
GRU	0.7	-	-	[38]
S-GAN	0.56	-	-	[38]
Hybrid (TDNN + ARIMA)	3.35	-	-	[49]
LSTM	0.14	3.04%	98.2%	Our model

* Best result

The models presented in Table 3 showcase various methods, including hybrid approaches, recurrent neural networks (RNNs), and generative models. The first model, a hybrid of ARIMA and LSTM, achieves an RMSE of 0.15 and a MAPE of 4.3%. However, the R² value is missing, making a comprehensive evaluation challenging. The second model, a stacked LSTM, performs similarly in terms of RMSE (0.15) but significantly better in MAPE (0.079%) and R² (96.8%), highlighting its superiority over the hybrid model. The third model based on gated recurrent unit (GRU) presents an RMSE of 0.7 without reporting the MAPE and R² values, and this makes the overall performance of the model difficult to be assessed. The fourth model uses sentiment analysis with a generative adversarial network (S-GAN) and achieves an RMSE of 0.56, but again, the lack of MAPE and R² values limits its evaluation. The fifth model, a hybrid of time delays neural network (TDNN) and ARIMA perform poorly with an RMSE of 3.35, with MAPE and R² values not provided for further assessment. Our LSTM model presented herein delivers outstanding results with an RMSE of 0.14, a MAPE of 3.04%, and an impressive R² value of 98.2%. These findings demonstrate the LSTM model's superior predictive accuracy and its ability to capture underlying patterns effectively. In conclusion, the comparison of various predictive models highlights the superiority of the stacked LSTM and the LSTM model presented in this paper. Both models exhibit remarkable performance with low RMSE, low

MAPE, and high R^2 values, showcasing the efficacy of LSTM-based architectures for the predictive task. Notably, the LSTM model presented in this paper stands out due to its simple design and effective feature engineering techniques. Despite its simplicity, our model demonstrates outstanding predictive accuracy, surpassing even the stacked LSTM in some metrics. This is due to the emphasis on feature engineering that has likely contributed to the model's ability to capture essential patterns in the data, leading to superior predictions. Moreover, one of the key advantages of the LSTM model presented in this paper is its ability to achieve such high performance with relatively low computation power and training time. This is crucial in practical applications where computational resources and time are essential. The model's efficiency in training and inference makes it an attractive choice for real-time predictions and large-scale deployments.

4. CONCLUSION

The study presented in this paper focuses on predicting essential commodity prices using an LSTM model enhanced by feature engineering. Our results indicate that this approach yields competitive performance compared to other models found in the literature. The proposed model demonstrated superior performance, particularly 0.14, 3.04, and 98.2% in RMSE, MAPE, and R^2 respectively. The simplicity and computational efficiency of the proposed model make it a promising approach for commodity price forecasting, especially in scenarios where computational resources may be limited. Future work could explore several directions. For instance, future models could incorporate external factors into the price prediction model. This could be achieved by using natural language processing techniques to analyze news headlines and social media sentiment related to the commodities under consideration, or by integrating economic indicators and climate data. Further research could also explore the use of other deep learning architectures for commodity price prediction. Hybrid models combining the strengths of different architectures could be particularly promising.

ACKNOWLEDGEMENTS

The authors would like to thank the Arab American University and the editors and reviewers for their comments and constructive feedback on the manuscript.

REFERENCES




- [1] Y. A. Shiferaw, "Time-varying correlation between agricultural commodity and energy price dynamics with Bayesian multivariate DCC-GARCH models," *Physica A: Statistical Mechanics and its Applications*, vol. 526, p. 120807, Jul. 2019, doi: 10.1016/j.physa.2019.04.043.
- [2] C. B. Barrett, "Heteroscedastic price forecasting for food security management in developing countries," *Oxford Development Studies*, vol. 25, no. 2, pp. 225–236, Jun. 1997, doi: 10.1080/13600819708424131.
- [3] G. Gereffi, "Global value chains in a post-Washington Consensus world," *Review of International Political Economy*, vol. 21, no. 1, pp. 9–37, Jan. 2014, doi: 10.1080/09692290.2012.756414.
- [4] P. Giordani and P. Söderlind, "Inflation forecast uncertainty," *European Economic Review*, vol. 47, no. 6, pp. 1037–1059, Dec. 2003, doi: 10.1016/S0014-2921(02)00236-2.
- [5] S. Umesh, H. M. G. Manukumar, and B. Chandrasekhar, "Sustainable Agriculture and Food Security," in *Biotechnology for Sustainable Agriculture*, Elsevier, 2018, pp. 67–92. doi: 10.1016/B978-0-12-812160-3.00003-9.
- [6] R. Ihle, M. K. El-Jafari, and S. von Cramon-Taubadel, "Effects of political instability on the volatility of Palestinian food prices," *New Medit*, vol. 18, no. 3, pp. 59–76, Sep. 2019, doi: 10.30682/nm1903e.
- [7] H. Ahumada and M. Cornejo, "Forecasting food prices: The case of corn, soybeans and wheat," *International Journal of Forecasting*, vol. 32, no. 3, pp. 838–848, Jul. 2016, doi: 10.1016/j.ijforecast.2016.01.002.
- [8] A. Aisen and F. J. Veiga, "Political instability and inflation volatility," *Public Choice*, vol. 135, no. 3–4, pp. 207–223, Jun. 2008, doi: 10.1007/s11127-007-9254-x.
- [9] H. Onyeaka, C. K. Anumudu, Z. T. Al-Sharify, E. Egele-Godswill, and P. Mbaegbu, "COVID-19 pandemic: A review of the global lockdown and its far-reaching effects," *Science Progress*, vol. 104, no. 2, p. 003685042110198, Apr. 2021, doi: 10.1177/00368504211019854.
- [10] R. J. Myers, "On the costs of food price fluctuations in low-income countries," *Food Policy*, vol. 31, no. 4, pp. 288–301, Aug. 2006, doi: 10.1016/j.foodpol.2006.03.005.
- [11] M. Uyttendaele, E. Franz, and O. Schlüter, "Food Safety, a Global Challenge," *International Journal of Environmental Research and Public Health*, vol. 13, no. 1, p. 67, Dec. 2015, doi: 10.3390/ijerph13010067.
- [12] M. Kalkuhl, J. von Braun, and M. Torero, "Volatile and Extreme Food Prices, Food Security, and Policy: An Overview," in *Food Price Volatility and Its Implications for Food Security and Policy*, Cham: Springer International Publishing, 2016, pp. 3–31. doi: 10.1007/978-3-319-28201-5_1.
- [13] M. F. Bellemare, "Rising Food Prices, Food Price Volatility, and Social Unrest," *American Journal of Agricultural Economics*, vol. 97, no. 1, pp. 1–21, Jan. 2015, doi: 10.1093/ajae/aau038.
- [14] J.-S. Chou and T.-K. Nguyen, "Forward Forecast of Stock Price Using Sliding-Window Metaheuristic-Optimized Machine-Learning Regression," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3132–3142, Jul. 2018, doi: 10.1109/TII.2018.2794389.
- [15] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques,"

- Procedia Computer Science*, vol. 167, pp. 599–606, 2020, doi: 10.1016/j.procs.2020.03.326.
- [16] A. Kuvalekar, S. Manchewar, S. Mahadik, and S. Jawale, "House Price Forecasting Using Machine Learning," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3565512.
 - [17] K. Noor and S. Jan, "Vehicle Price Prediction System using Machine Learning Techniques," *International Journal of Computer Applications*, vol. 167, no. 9, pp. 27–31, Jun. 2017, doi: 10.5120/ijca2017914373.
 - [18] A. M. Khedr, I. Arif, P. R. P. V. M. El-Bannany, S. M. Alhashmi, and M. Sreedharan, "Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey," *Intelligent Systems in Accounting, Finance and Management*, vol. 28, no. 1, pp. 3–34, Jan. 2021, doi: 10.1002/isaf.1488.
 - [19] L. Menculini *et al.*, "Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices," *Forecasting*, vol. 3, no. 3, pp. 644–662, Sep. 2021, doi: 10.3390/forecast3030040.
 - [20] Y. Ning, H. Kazemi, and P. Tahmasebi, "A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet," *Computers & Geosciences*, vol. 164, p. 105126, Jul. 2022, doi: 10.1016/j.cageo.2022.105126.
 - [21] J. Wang, G. Athanasopoulos, R. J. Hyndman, and S. Wang, "Crude oil price forecasting based on internet concern using an extreme learning machine," *International Journal of Forecasting*, vol. 34, no. 4, pp. 665–677, Oct. 2018, doi: 10.1016/j.ijforecast.2018.03.009.
 - [22] L. Yu, W. Dai, and L. Tang, "A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting," *Engineering Applications of Artificial Intelligence*, vol. 47, pp. 110–121, Jan. 2016, doi: 10.1016/j.engappai.2015.04.016.
 - [23] M. Långkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognition Letters*, vol. 42, pp. 11–24, Jun. 2014, doi: 10.1016/j.patrec.2014.01.008.
 - [24] A. R. S. Parmezan, V. M. A. Souza, and G. E. A. P. A. Batista, "Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model," *Information Sciences*, vol. 484, pp. 302–337, May 2019, doi: 10.1016/j.ins.2019.01.076.
 - [25] T. H. H. Aldhyani and A. Alzahrani, "Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms," *Electronics*, vol. 11, no. 19, p. 3149, Sep. 2022, doi: 10.3390/electronics11193149.
 - [26] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms," *AI*, vol. 2, no. 4, pp. 477–496, Oct. 2021, doi: 10.3390/ai2040030.
 - [27] Y. E. Shao and J.-T. Dai, "Integrated Feature Selection of ARIMA with Computational Intelligence Approaches for Food Crop Price Prediction," *Complexity*, vol. 2018, pp. 1–17, Jul. 2018, doi: 10.1155/2018/1910520.
 - [28] R. W. Divisekara, G. J. M. S. R. Jayasinghe, and K. W. S. N. Kumari, "Forecasting the red lentils commodity market price using SARIMA models," *SN Business & Economics*, vol. 1, no. 1, p. 20, Jan. 2021, doi: 10.1007/s43546-020-00020-x.
 - [29] R. Samsudin, A. Shabri, and P. Saad, "A Comparison of Time Series Forecasting using Support Vector Machine and Artificial Neural Network Model," *Journal of Applied Sciences*, vol. 10, no. 11, pp. 950–958, May 2010, doi: 10.3923/jas.2010.950.958.
 - [30] C. Z. Yuan and S. K. Ling, "Long Short-Term Memory Model Based Agriculture Commodity Price Prediction Application," in *Proceedings of the 2020 2nd International Conference on Information Technology and Computer Communications*, Aug. 2020, pp. 43–49, doi: 10.1145/3417473.3417481.
 - [31] M. Abumohsen, A. Y. Owda, and M. Owda, "Electrical Load Forecasting Using LSTM, GRU, and RNN Algorithms," *Energies*, vol. 16, no. 5, p. 2283, Feb. 2023, doi: 10.3390/en16052283.
 - [32] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
 - [33] S. Siarni-Namini, N. Tavakoli, and A. Siarni Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2018, pp. 1394–1401, doi: 10.1109/ICMLA.2018.00227.
 - [34] R. Ly, F. Traore, and K. Dia, "Forecasting commodity prices using long-short-term memory neural networks," 2021, doi: 10.2499/p15738coll2.134265.
 - [35] A. H. Krishnan, R. Murugesan, and E. Mishra, "Forecasting agricultural commodities prices using deep learning-based models: basic LSTM, bi-LSTM, stacked LSTM, CNN LSTM, and convolutional LSTM," *International Journal of Sustainable Agricultural Management and Informatics*, vol. 8, no. 3, p. 1, 2022, doi: 10.1504/IJSAMI.2022.10048228.
 - [36] Vasvani A. *et al.*, "Attention is all you need," *Advances in neural information processing systems*, pp. 5998–6008, 2017, doi: 10.48550/arXiv.1706.03762.
 - [37] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>
 - [38] P. Sonkiya, V. Bajpai, and A. Bansal, "Stock price prediction using BERT and GAN," 2021, [Online]. Available: <http://arxiv.org/abs/2107.09055>
 - [39] M. Gaur, K. Faldu, and A. Sheth, "Semantics of the Black-Box: Can Knowledge Graphs Help Make Deep Learning Systems More Interpretable and Explainable?," *IEEE Internet Computing*, vol. 25, no. 1, pp. 51–59, Jan. 2021, doi: 10.1109/MIC.2020.3031769.
 - [40] M. A. Hedderich and D. Klakow, "Training a Neural Network in a Low-Resource Setting on Automatically Annotated Noisy Data," in *Proceedings of the Workshop on Deep Learning Approaches for Low-Resource NLP*, 2018, pp. 12–18, doi: 10.18653/v1/W18-3402.
 - [41] WFP, "World Food Programme." <https://www.wfp.org/countries/palestine>
 - [42] B. Ghosh, B. Basu, and M. O'Mahony, "Multivariate Short-Term Traffic Flow Forecasting Using Time-Series Analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 246–254, Jun. 2009, doi: 10.1109/TITS.2009.2021448.
 - [43] G. P. Zhang and M. Qi, "Neural network forecasting for seasonal and trend time series," *European Journal of Operational Research*, vol. 160, no. 2, pp. 501–514, Jan. 2005, doi: 10.1016/j.ejor.2003.08.037.
 - [44] R. Reghunath, T. R. S. Murthy, and B. R. Raghavan, "Time Series Analysis to Monitor and Assess Water Resources: A Moving Average Approach," *Environmental Monitoring and Assessment*, vol. 109, no. 1–3, pp. 65–72, Oct. 2005, doi: 10.1007/s10661-005-5838-4.
 - [45] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, Jan. 2004, doi: 10.1016/j.ijforecast.2003.09.015.
 - [46] J. T. Scruggs and F. Nardari, "Why Does Stock Market Volatility Change Over Time? A Time-Varying Variance Decomposition for Stock Returns," *SSRN Electronic Journal*, 2005, doi: 10.2139/ssrn.676673.
 - [47] M. A. Gruet, M. Chandorkar, A. Sicard, and E. Camporeale, "Multiple-Hour-Ahead Forecast of the Dst Index Using a Combination of Long Short-Term Memory Neural Network and Gaussian Process," *Space Weather*, vol. 16, no. 11, pp. 1882–1896, Nov. 2018, doi: 10.1029/2018SW001898.




- [48] G. Morse and K. O. Stanley, "Simple Evolutionary Optimization Can Rival Stochastic Gradient Descent in Neural Networks," in *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, Jul. 2016, pp. 477–484. doi: 10.1145/2908812.2908916.
- [49] K. Girish and S. Kanchan, "Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System," *Agricultural Economics Research Review*, vol. 26, no. 2, pp. 229–239, 2013, doi: 10.22004/ag.econ.162150.

BIOGRAPHIES OF AUTHORS



Mohammad Tami    is a senior software engineer at Apple. Mr. Tami always had a passion for technology and computer engineering. After completing his undergraduate degree at A Najah National University, Tami began his career working with various US-based companies in web-based applications and data-intensive systems. Tami's focus is on developing innovative solutions for data-intensive applications using advanced techniques in ML and AI. As he gained more experience in the field, he decided to further his education by pursuing a master's degree in data science at the Arab American University. As a dedicated and talented computer engineer, Mr. Tami constantly seeks new challenges and opportunities to further his knowledge and skills. He can be contacted using the following emails: m.abutami@student.aaup.edu or mabutame@gmail.com.



Amani Yousef Owda    assistant professor in Computer Engineering and Data Science in the Faculty of Graduate Studies at the Arab American University in Palestine. She worked as a research associate in the Faculty of Engineering at the University of Manchester from 2019 -2020. In addition, she worked in the School of Engineering at Manchester Metropolitan University from 2015 - 2019. She worked at Birzeit University from 2007- 2011. She received her MSc. degree (Hons.) from The University of Manchester, UK in 2013, and her Ph.D. degree in computer engineering from Manchester Metropolitan University, UK in 2018. She is currently an assistant professor in the Department of Natural, Engineering, and Technology Sciences, at the Arab American University in Palestine. Since 2018, she leads research in multi-disciplinary fields with a focus on artificial intelligence, machine learning, decision support systems, image processing, medical applications of microwave and millimeter-wave imaging, security screening, and anomaly detection. She has published more than 42 articles in well reputable journals. She is a reviewer in many well-known Journals, and she is supervising MSc and PhD students. She can be contacted using the following emails: Amani.Owda@aaup.edu or amaniabubaha@gmail.com.