International Journal of Artificial Intelligent and Informatics
Vol. 3, No. 1, January 2025, pp. 9~ 15
ISSN: 2622-626X, DOI: http://doi.org/10.33292/ijarlit.v3i1.41

Performance Comparison of Long Short-Term Memory and Convolutional Neural Network for Prediction of Exchange Rate of Indian Rupee against US Dollar

Kovat Rai[1*], Amit Vijayan[1]
[1] Birla Institute of Technology and Science, India
Email: kovatrai@gmail.com

Article Info

Article history:

Received: July 20, 2024
Revised: September 30, 2024
Accepted: December 10, 2024
Available Online: January 30, 2025

Keywords:

exchange rate prediction
Indian Rupee
US Dollar
deep learning
LSTM
CNN
financial time series

ABSTRACT

This study compares the effectiveness of Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models in predicting the exchange rate of the Indian Rupee (INR) against the United States Dollar (USD). Using historical data from 2017 to 2023 obtained from Yahoo Finance, both models were trained and evaluated based on several performance metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and accuracy. The results showed that the hybrid LSTM model consistently outperformed the CNN model on all evaluation metrics, with a Test RMSE value of 0.38 compared to 1.32 for CNN. The LSTM model also showed better stability between training and testing performance, indicating better generalization ability and lower risk of overfitting. These findings confirm the superiority of the LSTM architecture in capturing the complex temporal patterns inherent in financial time series data, making it a more reliable option for currency exchange rate prediction.

1. INTRODUCTION

Currency exchange rates are one of the most important economic indicators as they have a significant influence on various aspects of the economy such as international trade, inflation, foreign investment, and overall economic stability. Predicting currency exchange rate movements with high accuracy has substantial strategic value for various stakeholders including governments, central banks, international businesses, and investors .[1], [2]

The Indian Rupee (INR) as the currency of one of the world's largest economies plays an important role in the global economy. Fluctuations in the INR exchange rate against the United States Dollar (USD) have far-reaching implications for the Indian economy and its trading partners. Therefore, the development of accurate prediction models for the INR/USD exchange rate is an important focus in finance and economics research .[3]

In recent years, advances in the field of artificial intelligence and machine learning have opened up new opportunities in developing more sophisticated financial prediction models. In particular, deep learning techniques such as Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) have shown significant potential in modeling and predicting complex and non-stationary financial time series .[4], [5]

LSTM, which is a variant of Recurrent Neural Network (RNN), has proven to be effective in handling long-term dependencies often present in financial time series data. The LSTM's ability to "remember" relevant information and "forget" irrelevant information through its special gating mechanism makes it well suited for modeling complex temporal sequences such as currency exchange rates .[6], [7]

Previous research on the use of LSTM (Long Short-Term Memory) for time series forecasting has shown significant progress in overcoming the limitations of traditional forecasting methods. LSTM, as a type of Recurrent Neural Network (RNN), is able to handle long-term dependencies and non-linearities in time series data, so it is widely chosen in various applications. For example, Bouktif et al. Bouktif et al.[8] developed an optimal LSTM model for electricity load forecasting by using feature selection and genetic algorithm, where this model showed more stable performance in handling data fluctuations than shallow neural network approaches. Similarly, Sood et al.[9] compared the performance of ARIMA model with LSTM in ozone concentration forecasting, and the results confirmed the superiority of LSTM in capturing non-linear patterns and handling complex data.

In addition, the application of LSTM has also been explored in various time series forecasting application domains. In USD/CNY exchange rate forecasting, Cao et al. Cao et al. (2020) proposed a deep coupled LSTM approach that is able to capture the interrelationships between various economic variables. Meanwhile, Zaheer et al.[10] and Yıldırım et al.[11] applied LSTM for forecasting stock and forex price movements by combining technical and macroeconomic indicators, which showed improved accuracy in predicting the direction of market movements. Not only limited to the financial sector, LSTM is also proven effective in the context of health and epidemiology, as Rashed and Hirata[12] did to predict positive cases of COVID-19 by integrating meteorological data and community mobility.

Other research has focused on improving the LSTM model through the incorporation of advanced techniques. Li and Bastos[13] in a systematic review study confirmed that the LSTM technique dominates stock market forecasting applications due to its flexible ability to model market dynamics. Pranolo et al.[14] developed a more robust LSTM model with parameter settings using PSO and bifold-attention mechanisms, resulting in adaptive performance on multivariate datasets with daily, weekly, and monthly time resolution. In addition, Sagheer and Kotb[15] proposed an unsupervised pre-training technique on the LSTM-based stacked autoencoder, which enables improved convergence and forecasting accuracy on multivariate time series datasets.

On the other hand, CNN, which was originally developed for image processing, has been adapted for time series analysis with its ability to detect local patterns through convolutional filters. Several studies have demonstrated the effectiveness of CNNs in extracting features and patterns from financial time series data, including stock prices and currency exchange rates .[16], [17]

Previous research on Convolutional Neural Network (CNN) for time series forecasting shows a diversity of approaches in integrating spatial and temporal features to improve prediction accuracy in various domains. CNN, as one of the deep learning methods, is known to be effective in extracting non-linear features from complex time series data, especially when combined with other architectures such as Long Short-Term Memory (LSTM) or Bidirectional LSTM (BiLSTM) to capture temporal dynamics. For example, Hao et al.[18] proposed a hybrid model that utilizes CNN to extract spatial features as well as LSTM to capture temporal dependencies in wind power forecasting. This approach proved to be effective in processing large-scale data and producing accurate predictions on time series data, especially for energy applications.

While both approaches have been applied separately in various financial prediction studies, direct comparisons between LSTMs and CNNs in the context of INR/USD exchange rate prediction are limited. Existing studies often do not provide a comprehensive evaluation based on various performance metrics or use limited data over shorter time spans.

To fill this research gap, our study aims to conduct a systematic comparison between LSTM and CNN hybrid models in predicting the INR/USD exchange rate using comprehensive historical data from 2017 to 2023. This study specifically aims to develop and optimize a hybrid model that integrates Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) to predict INR/USD exchange rate movements. In addition, this study is designed to conduct a comprehensive evaluation by comparing the performance of both models through various statistical metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and prediction accuracy. Furthermore, this study will deeply analyze the advantages and limitations of each approach in the context of currency exchange rate prediction, considering aspects such as the ability to capture temporal patterns, robustness to data noise, and computational complexity. Based on the empirical findings obtained, this study is expected to provide optimal methodological recommendations regarding the INR/USD exchange rate prediction approach, which can serve as a reference for researchers and practitioners in the field of economics and finance.

The results of this study are expected to provide valuable insights for academics, financial practitioners, and policy makers in developing and implementing more accurate and reliable currency exchange rate prediction models.

2. METHOD

2.1. Data Sources and Preparation

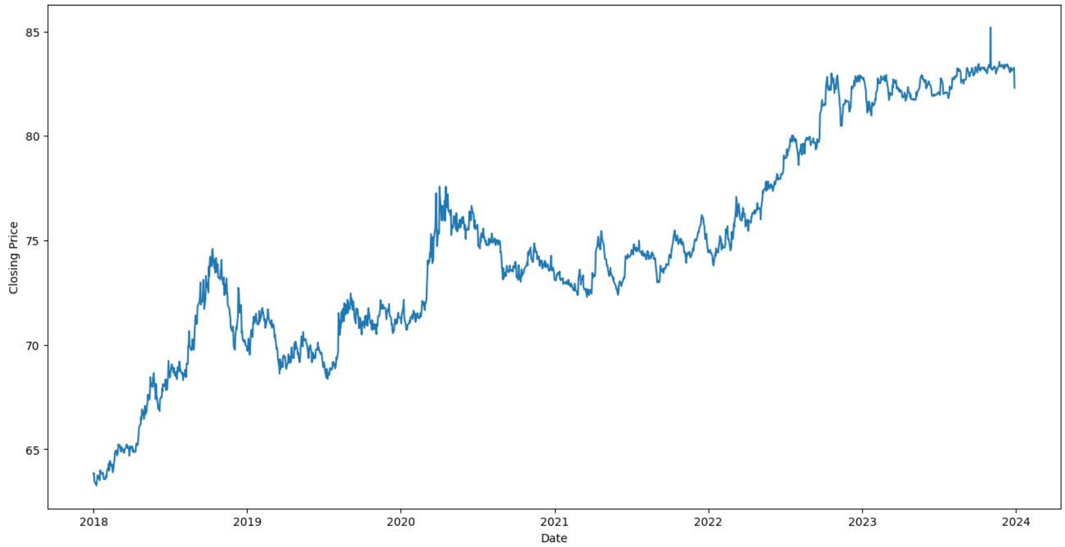

Figure 1. Exchange rate movement of Indian Rupee against USD

The dataset used in this study consists of historical data of the daily exchange rate of the Indian Rupee (INR) against the United States Dollar (USD) for the period January 1, 2017 to December 31, 2023. The data was obtained from Yahoo Finance, a source that has been validated in numerous previous financial studies[7] . The dataset includes information on the opening price (open), highest price (high), lowest price (low), closing price (close), and daily trading volume.

In the pre-processing stage, several steps are taken to ensure data quality and suitability:

1. Data cleaning: Identification and handling of missing values and outliers using linear interpolation and box-plot analysis.
2. Normalization: The data is normalized using the Min-Max Scaling technique to transform the values into the range [0,1]. This normalization is important to improve the convergence and performance of the machine learning model .[19]
3. Data Division: The dataset is divided into training set (70%) and testing set (30%) by maintaining the chronological order to ensure the temporal integrity of the data.
4. Feature Construction: Additional features are created through feature engineering techniques, including technical indicators such as Moving Average (MA), Relative Strength Index (RSI), and Bollinger Bands, as well as lag features of previous exchange rates to capture temporal dependencies.

2.2. Model Architecture

2.2.1. LSTM Model

The hybrid LSTM model developed in this study combines the LSTM layer with several other architectural components to improve predictive capabilities. The model architecture consists of:

1. Input Layer: Receives an n-dimensional feature vector, where n is the number of features after feature engineering.
2. LSTM layers: Two consecutive LSTM layers with 128 and 64 units respectively. The first layer returns the complete sequence for input to the second layer, while the second layer only returns the final output.
3. Dropout Layer: Applied after each LSTM layer with a dropout rate of 0.2 to prevent overfitting .[20]
4. Dense layers: Two fully connected layers with 32 and 16 neurons, respectively, with ReLU activation to introduce non-linearity.
5. Output Layer: Single dense layer with linear activation for exchange rate regression.

2.2.1. CNN Model

The CNN model for time series prediction is designed to capture local patterns in exchange rate data. The model architecture consists of:

1. Input Layer: Receive reformatted time series data as 2D input with dimensions (sequence_length, n_features).
2. Convolutional Layers: Three consecutive 1D convolutional layers with 64, 32, and 16 filters, a kernel size of 3, and ReLU activation.

3. Max Pooling Layer: After each convolutional layer to reduce dimensionality and capture the most salient features.
4. Flatten Layer: To convert the convolutional output into a 1D vector before input to the dense layer.
5. Dense layer: Two fully connected layers with 32 and 16 neurons, respectively, with ReLU activation.
6. Output Layer: Single dense layer with linear activation for exchange rate regression.

2.3. Model Training

Both models were trained with similar training configurations to ensure a fair comparison:

1. Optimization Algorithm: Adam optimizer with an initial learning rate of 0.001 and decreasing adaptive learning rate.
2. Loss function: Mean Squared Error (MSE) to measure the difference between predicted and actual values.
3. Batch Size: 32 to find a balance between training speed and accuracy.
4. Epochs: Maximum 100 epochs with implementation of early stopping with patience 10 to prevent overfitting.
5. Validation: 20% of the training data is used as a validation set to monitor performance during training.

Additional callbacks including model checkpointing are implemented to save the best performing model based on validation loss.

2.4. Model Evaluation

The performance of both models was evaluated using several metrics to provide a comprehensive perspective:

1. Root Mean Square Error (RMSE): Measures the square root of the average square of the difference between the predicted value and the actual value.
2. Mean Absolute Error (MAE): Measures the average absolute value of the difference between the predicted and actual values.
3. Mean Absolute Percentage Error (MAPE): Measures the average percentage absolute error relative to the true value.
4. Coefficient of Determination (R^2): Measures the proportion of variation in the dependent variable that can be explained by the independent variable.
5. Accuracy: Calculated as 100% - MAPE, shows the level of model accuracy in percentage form.

The evaluation was conducted on both data sets (training and testing) to understand the model's ability to handle both pre-seen and unseen data.

3. RESULTS AND DISCUSSION

3.1. Model Performance Comparison

The performance evaluation results of both models on training and testing data are presented in Table 1. Evaluation metrics include Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and accuracy.

Table 1: Comparison of evaluation metrics of hybrid LSTM and CNN models

Metrics	LSTM (Train)	LSTM (Test)	CNN (Train)	CNN (Test)
RMSE	0,48	0,38	1,13	1,32
MAE	0,37	0,30	0,91	1,25
MAPE	0,51%	0,36%	1,25%	1,52%
R^2	0,9781	0,6142	0,8866	-3,2203
Accuracy	99,49%	99,64%	98,75%	98,48%

Based on the evaluation results presented in Table 1, the comparative analysis shows significant performance differences between the hybrid LSTM and CNN models in predicting the INR/USD exchange rate. The hybrid LSTM model consistently shows superiority in terms of prediction accuracy, with the RMSE value on the test data of 0.38 being significantly lower than that of the CNN model (1.32), indicating a significantly smaller error rate. The generalization ability of the LSTM model also proved superior, indicated by the minimal performance difference between the training and testing data, while the CNN model experienced a marked increase in error on the testing data, hinting at an overfitting problem. The coefficient of determination (R^2) analysis revealed a more striking difference - although both models exhibited high R^2 values on the training data (0.9781 for LSTM and 0.8866 for CNN), the LSTM model maintained a positive value (0.6142) on the testing data, while the CNN model produced a strongly negative value (-3.2203) indicating an inability to capture underlying trends in the new data. From the perspective of percentage error, the LSTM model showed superiority with a MAPE of only 0.36% on the test data, versus 1.52% on the CNN model, which makes it more suitable for financial applications that require high precision. In addition, the LSTM model displays better performance consistency between training and testing data, with some metrics such as MAPE and accuracy even showing improvement on testing data, reinforcing its position as a more reliable approach under various market conditions. These findings collectively confirm the superiority of the hybrid LSTM model in the context of currency exchange rate prediction.

3.2. Visual Analysis of Prediction Results

A visual analysis of the predictions of both models compared to the actual values on the test dataset is presented in Figure 2 and Figure 3. These visualizations help in understanding how both models capture trends and fluctuations in the INR/USD exchange rate.

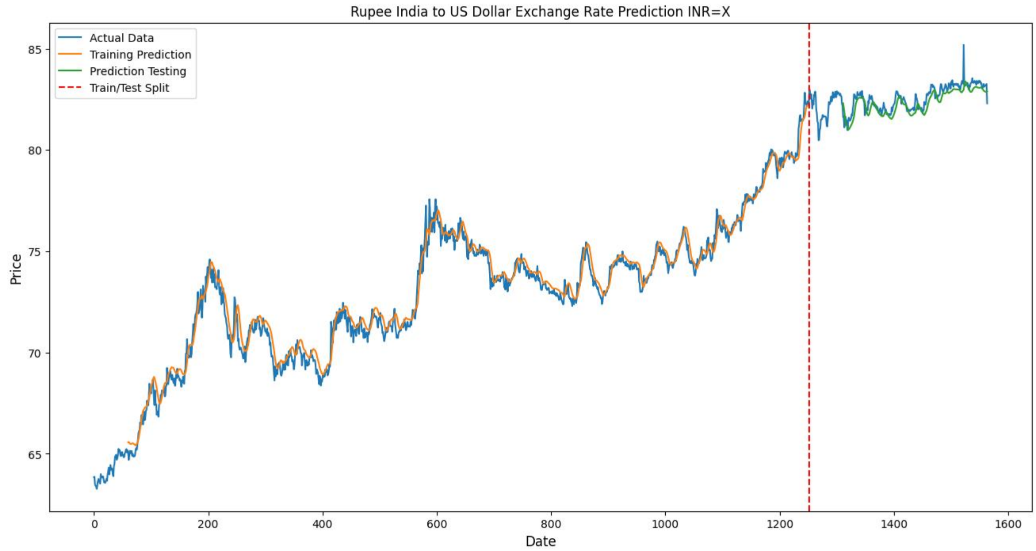

Figure 2. Prediction graphs of both LSTM Models

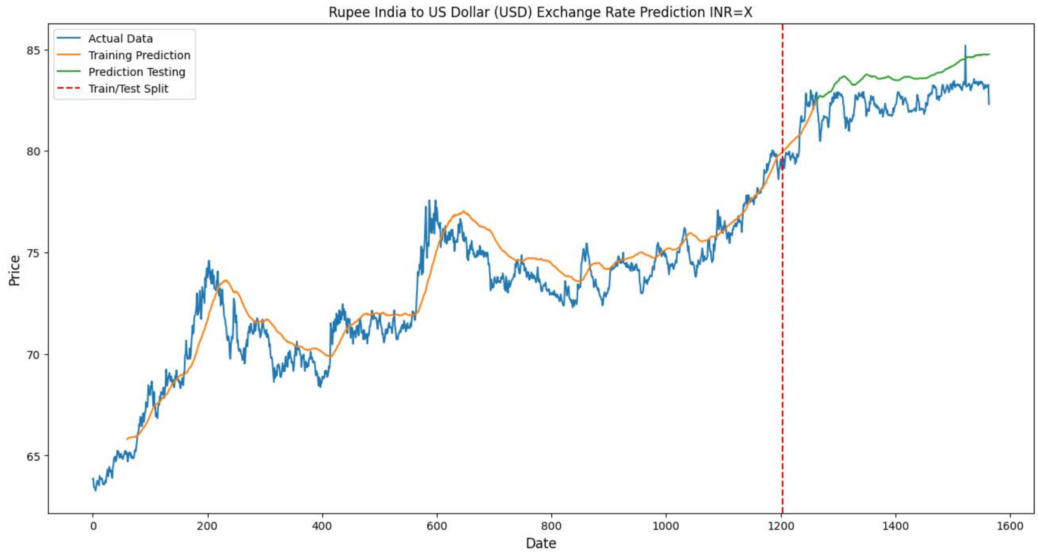

Figure 3 Prediction graphs of both CNN Models

Figure 2 shows that the LSTM model is able to follow the general trend and fluctuations in the exchange rate very well, with minimal deviations from the actual values. The model successfully captures both short-term and long-term movements in the data.

In contrast, as seen in Figure 3, CNN models tend to produce predictions that deviate more from actual values and often fail to capture directional changes and significant fluctuations in exchange rates. The model appears to struggle particularly during periods of high volatility, where the prediction error becomes larger.

3.3. Discussion of Model Performance

The superior performance of the hybrid LSTM model in this study can be explained by several interrelated factors.

First, the architecture suitable for temporal data is the main advantage of LSTM. The recurrent structure of LSTM that is specifically designed to handle sequence dependencies and long-term memory is very suitable for financial time series data. The gate mechanism in LSTM (input gate, forget gate, and output gate) allows the model to selectively retain or discard information from previous time steps, which is essential for modeling complex foreign exchange market dynamics.

Secondly, the ability to capture irregular patterns is an added value of LSTM. Currency exchange rates are often influenced by various factors such as monetary policy, geopolitical events, and market sentiment which can lead to irregular and non-stationary patterns. LSTM has proven to excel in capturing these kinds of patterns compared to traditional models and some other deep learning architectures.

Third, the suitability of the data structure provides an advantage to LSTM. While CNNs are very effective for detecting spatial patterns in data (such as in image processing applications), their ability to capture long-term temporal dependencies may be limited compared to LSTMs, despite using 1D convolutions designed for sequential data.

Fourth, the overfitting problem in CNN is a significant drawback. The highly negative R^2 value on the test data for the CNN model indicates a serious overfitting problem, where the model may be too specifically modeling the noise in the training data rather than the underlying trend. Although regularization techniques such as dropout and max pooling were applied, the CNN model still showed difficulty in generalization to new data.

Fifth, an appropriate balance of model complexity is achieved by LSTM. Hybrid LSTM models with a relatively simple architecture (two LSTM layers followed by several dense layers) seem to achieve a good balance between model complexity and generalization ability, while CNNs may experience problems with too many parameters relative to the pattern structure in the data.

Finally, better prediction stability was demonstrated by the LSTM model. The small difference between training and testing performance for the LSTM model indicates better stability, which is an important characteristic for financial prediction models to be used in real-world decision making.

3.3. Practical Implications

The findings of this study have several practical implications for currency exchange rate prediction and the application of deep learning techniques in financial analysis:

1. Model Selection: For currency exchange rate prediction applications, LSTM-based architectures seem to be more recommended compared to CNNs based on comprehensive performance metrics.
2. Trading System Development: The high level of accuracy demonstrated by the LSTM model (99.64% on test data) shows the potential for the development of automated trading systems that can generate profits by exploiting accurate predictions of exchange rate movements.
3. Risk Management: More accurate predictions from the LSTM model can assist financial institutions and multinational corporations in more effective exchange rate risk management, potentially resulting in significant savings.
4. Hybrid Model Development: Although CNNs independently do not show good performance, the combination of CNNs with LSTMs in a more complex hybrid architecture (CNN-LSTM) could possibly result in further performance improvements, as suggested by some recent research .[21]

4. CONCLUSION

This study conducts a systematic comparison between hybrid LSTM and CNN models in predicting the exchange rate of the Indian Rupee (INR) against the United States Dollar (USD) using historical data from 2017 to 2023. The comprehensive evaluation results show that the hybrid LSTM model consistently outperforms CNN in all evaluation metrics, with significant advantages in prediction accuracy and generalization ability. The most striking performance difference was seen in the test R^2 values, where LSTM maintained a solid positive value (0.6142), while CNN produced a strongly negative value (-3.2203), indicating the CNN model's inability to capture basic trends in the new data. Although both models achieve high prediction accuracy (>98%), LSTM exhibits a much lower error rate (RMSE 0.38 vs. 1.32; MAPE 0.36% vs. 1.52%), a critical factor in practical applications such as algorithmic trading and foreign exchange risk management. This advantage of LSTM is mainly due to its recurrent architecture that effectively captures long-term temporal dependencies, an essential characteristic in financial time series data. In addition, LSTM shows resilience to overfitting with minimal gap between training and testing performance, confirming its better generalization capacity compared to CNN.

The findings make an important contribution to the exchange rate prediction literature by reinforcing the position of LSTM as a superior architecture for financial time series modeling, while exposing the limitations of CNN in this domain. However, this research has several limitations that open up opportunities for further studies, including the exploration of a hybrid CNN-LSTM architecture to combine the strengths of both approaches, testing the robustness of the model under various market conditions (including periods of extreme volatility), as well as the integration of exogenous variables such as macroeconomic indicators and market sentiment data. Future research is also recommended to develop and test model-based trading strategies to evaluate their economic advantages in practice, as well as compare this approach with traditional methods to provide a more comprehensive perspective on the effectiveness of deep learning in exchange rate prediction.

REFERENCES

[1] F. Shen, J. Chao, and J. Zhao, "Forecasting exchange rate using deep belief networks and conjugate gradient method," *Neurocomputing*, vol. 167, pp. 243-253, 2015.

[2] S. Galeshchuk and S. Mukherjee, "Deep networks for predicting direction of change in foreign exchange rates," *Intell. Syst. Accounting, Financ. Manag.*, vol. 24, no. 4, pp. 100-110, 2017.

[3] M. Kumar and M. Thenmozhi, "Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models," *Int. J. Banking, Account. Financ.*, vol. 5, no. 3, pp. 284-308, 2014.

[4] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654-669, 2018.

[5] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005-2019," *Appl. Soft Comput.*, vol. 90, p. 106181, 2020.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.

[7] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, 2018, pp. 1394-1401.

[8] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting Using Feature Selection and Genetic Algorithm: Comparison With Machine Learning Approaches†," *Energies*, vol. 11, no. 7, p. 1636, 2018, doi: 10.3390/en11071636.

[9] H. Sood, D. Narula, and P. S. Rana, "Forecasting Ozone Concentration Data: Arima v/S LSTM," *Int. J. Eng. Appl. Sci. Technol.*, vol. 04, no. 04, pp. 378-382, 2019, doi: 10.33564/ijeast.2019.v04i04.060.

[10] S. Zaheer *et al.*, "A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model," *Mathematics*, vol. 11, no. 3, p. 590, 2023, doi: 10.3390/math11030590.

[11] D. C. Yıldırım, İ. H. Toroslu, and U. Fiore, "Forecasting Directional Movement of Forex Data Using LSTM With Technical and Macroeconomic Indicators," *Financ. Innov.*, vol. 7, no. 1, 2021, doi: 10.1186/s40854-020-00220-2.

[12] E. A. Rashed and A. Hirata, "One-Year Lesson: Machine Learning Prediction of COVID-19 Positive Cases With Meteorological Data and Mobility Estimate in Japan," *Int. J. Environ. Res. Public Health*, vol. 18, no. 11, p. 5736, 2021, doi: 10.3390/ijerph18115736.

[13] A. W. Li and G. S. Bastos, "Stock Market Forecasting Using Deep Learning and Technical Analysis: A Systematic Review," *IEEE Access*, vol. 8, pp. 185232-185242, 2020, doi: 10.1109/ACCESS.2020.3030226.

[14] A. Pranolo, Y. Mao, A. P. Wibawa, A. B. P. Utama, and F. A. Dwiyanto, "Robust LSTM With Tuned-PSO and Bifold-Attention Mechanism for Analyzing Multivariate Time-Series," *IEEE Access*, vol. 10, no. July, pp. 78423-78434, 2022, doi: 10.1109/ACCESS.2022.3193643.

[15] A. Sagheer and M. Kotb, "Unsupervised Pre-Training of a Deep LSTM-based Stacked Autoencoder for Multivariate Time Series Forecasting Problems," *Sci. Rep.*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-55320-6.

[16] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *arXiv Prepr. arXiv1703.04691.*, 2017.

[17] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. neural networks Learn. Syst.*, vol. 30, no. 11, pp. 3212-3232, 2019.

[18] Z. Hao, D. Niu, M. Yu, K. Wang, Y. Liang, and X. Xu, "A Hybrid Deep Learning Model and Comparison for Wind Power Forecasting Considering Temporal-Spatial Feature Extraction," *Sustainability*, vol. 12, no. 22, p. 9490, 2020, doi: 10.3390/su12229490.

[19] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, "Stock market prediction based on generative adversarial network," *Procedia Comput. Sci.*, vol. 147, pp. 400-406, 2019.

[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929-1958, 2014.

[21] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, pp. 17351-17360, 2020.