

Forecasting Next-Time-Step Forex Market Stock Prices Using Neural Networks

Mahdi Navaei^{1*} and Mostafa Pahlevanzadeh²

¹Information technology graduate, University of Applied Science and Technology Informatics of Iran, Tehran, Iran

²Head of Computer Science Department, University of Applied Science and Technology Informatics of Iran, Tehran

Corresponding Author

Mahdi Navaei, Information Technology graduate, University of Applied Science and Technology Informatics of Iran, Tehran, Iran.

Submitted: 2024 Apr 15; Accepted: 2024 May 07; Published: 2024 May 16

Citation: Navaei, M., Pahlevanzadeh, M. (2024). Forecasting Next-Time-Step Forex Market Stock Prices Using Neural Networks. *Adv Mach Lear Art Inte*, 5(2), 01-10.

Abstract

Purpose: This study aims to predict the closing price of the EUR/JPY currency pair in the forex market using recurrent neural network (RNN) architectures, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), with the incorporation of Bidirectional layers.

Methods: The dataset comprises hourly price data obtained from Yahoo Finance and pre-processed accordingly. The data is divided into training and testing sets, and time series sequences are constructed for input into the models. The RNN, LSTM, and GRU models are trained using the Adam optimization algorithm with the mean squared error (MSE) loss metric.

Results: Results indicate that the LSTM model, particularly when coupled with Bidirectional layers, exhibits superior predictive performance compared to the other models, as evidenced by lower MSE values.

Conclusions: Therefore, the LSTM model with Bidirectional layers is the most effective in predicting the EUR/JPY currency pair's closing price in the forex market. These findings offer valuable insights for practitioners and researchers involved in financial market prediction and neural network modelling.

Keywords: Financial Market, Deep Learning, Time Series Analysis, Forex Trading, Neural Networks

1. Introduction

These days, artificial intelligence is being utilized across various industries, from predicting heart diseases to accurately analyzing stock markets, which is invaluable for market traders [1,2]. Time series financial forecasting is among the challenging and crucial tasks, and researchers strive to extract underlying patterns for predicting future market trends [3]. The primary challenge in this regard is enhancing the prediction accuracy and precision, which researchers aim to improve by presenting various models and methodologies.

However, due to the nature of volatile stock markets, models face limitations and challenges. Many researchers have focused on price prediction models and utilized criteria indicating the proximity of predicted prices to actual ones for evaluation. Another challenge in prediction models is the disregard for predicting the trend of stock

price movements [4].

This oversight can lead to erroneous decisions and losses for investors. For instance, a trader may enter a buying position based on a price increase prediction while the actual price decreases, resulting in losses. The reason for such occurrences lies in neglecting the trend of price changes in prediction. To overcome these challenges, after reviewing extensive research in this field, we propose a model in this article that utilizes collective learning from basic artificial intelligence models to maximize prediction results. Improving accuracy and prediction power in these markets can assist traders and investors in making better decisions regarding buying or selling currencies [5].

In this regard, the use of neural network algorithms and time series models has received considerable attention as powerful

tools in predicting forex markets. In this article, we examine the performance of three neural network algorithms, namely RNN (Recurrent Neural Network), LSTM (Long Short-Term Memory), and GRU (Gated Recurrent Unit), in forecasting the forex market, using the EUR/JPY currency pair dataset as an example. This dataset includes the price history of the euro and the yen in the forex market [6].

2. Research Methodology

In this study, we utilized three types of neural networks to predict the future EUR/JPY price at the subsequent step. The first model employed is the recurrent neural network (RNN), followed by the long short-term memory (LSTM) network, and lastly, the gated recurrent unit (GRU). The performance evaluation of each neural network is assessed using the Mean Squared Error (MSE) criterion.

2.1 Recurrent Neural Network (RNN)

The recurrent neural network (RNN) is a neural network architecture specifically designed to model temporal dependencies within sequential data.

Given the sequential or time-series nature of the data under consideration, conventional neural networks often fall short in effectiveness. Hence, it is more prudent to turn towards architectures capable of processing sequence-based data. Recurrent Neural Networks (RNNs), abbreviated as RNNs, possess the ability to discern patterns and glean insights from sequentially-ordered datasets [7,8].

In Recurrent Neural Networks (RNNs), the network architecture is engineered to incorporate a form of memory, allowing it to process a sequence of n data points sequentially [9].

As illustrated in Figure 1, at time t , the output h is fed into the subsequent block alongside the subsequent input x . Subsequently,

this block, by virtue of multiplying the inputs with the weights W_{hh} and W_{xh} , coupled with a bias term and, after that, applying an activation function and another weight, generates both the subsequent output and hidden state. These weight parameters and bias terms are dynamic elements that the algorithm adaptively learns during the training process.

Each output node, via its associated weight parameters and bias terms, learns from the training dataset. Fundamentally, an RNN behaves akin to simple neural networks, albeit with the added sophistication of its constituent blocks and their inherent memory capabilities.

During the training phase, akin to conventional neural networks, the weight parameters necessitate updates via backpropagation of errors. However, due to the temporal dependency inherent in RNNs, the error The backpropagation mechanism undergoes slight modifications, known as backpropagation through time (BPTT). A crucial aspect of the RNN algorithm is the presence of five trainable weights per block to facilitate updates during error backpropagation, thereby facilitating learning from the training dataset [10].

Furthermore, within RNNs, the input hidden state (h) assumes the form of a vector comprising n elements. The selection of the number of elements (n) is contingent upon user discretion during network instantiation, with higher values offering increased capability to discern intricate patterns [11]. Consequently, a small scale internal neural network is leveraged within RNNs for amalgamating inputs (x) and hidden states (h) to facilitate weight updates and learning. This algorithmic framework exhibits proficiency in detecting complex and dynamic patterns within data, rendering it suitable for deployment in financial market forecasting endeavours.

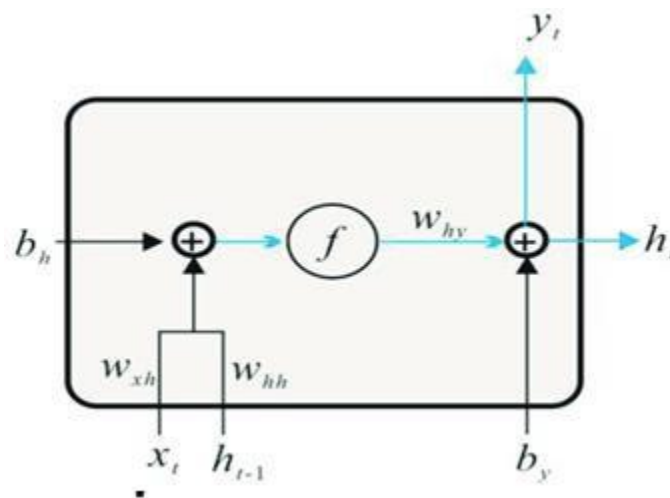


Figure 1: Conceptual Depiction of the Recurrent Neural Network (RNN) Architecture

2.2 The Long Short-Term Memory (LSTM) Neural Network

LSTM, a specialized variant of recurrent neural networks (RNNs), incorporates long short-term memory gates to retain and utilize past information in predictive tasks effectively. This model, belonging to the category of Recurrent Neural Networks (RNNs), is tailored for processing sequential data with prolonged dependencies [12]. Its memory management capability over time allows LSTM to discern intricate and protracted patterns within sequential datasets [13].

The rationale for employing the LSTM model in forecasting the forex market is as follows:

1) Long Short-Term Memory Capability: LSTM's distinctive memory units, known as cell states, enable the model to retain temporal information with extended dependencies. This feature facilitates leveraging long-term variations in the forex market, including trends, patterns, and market cycles, leading to notable enhancements in prediction accuracy [14].

2) Gate Management for Information Flow Control: LSTM proficiently regulates information flow within the network through three primary gates: the Forget Gate, Input Gate, and Output Gate. These gates empower the model to decide the retention, discarding, and output of information, thereby augmenting its predictive performance [15,16].

3) Processing Multi-dimensional Inputs: The LSTM model exhibits versatility in processing multidimensional inputs. In the context of forex market prediction, this attribute proves invaluable as the forex market encompasses diverse features such as market prices, trading volumes, technical indicators, and economic variables. LSTM adeptly accommodates these multi-dimensional features, utilizing them effectively in the prediction process [17].

Employing the LSTM model yields substantial improvements in forex market prediction. With its robust processing capabilities and adeptness in detecting intricate patterns, LSTM facilitates the identification of various price changes, trends, and market fluctuations, thereby furnishing accurate forecasts of future market behavior. Such forecasts can significantly benefit traders and investors, aiding them in making informed and timely decisions in the market [18,12].

In summary, the LSTM model, characterized by its long short-term memory capability, gate management mechanisms, and adeptness in processing multi-dimensional data, emerges as a potent tool for forex market prediction.

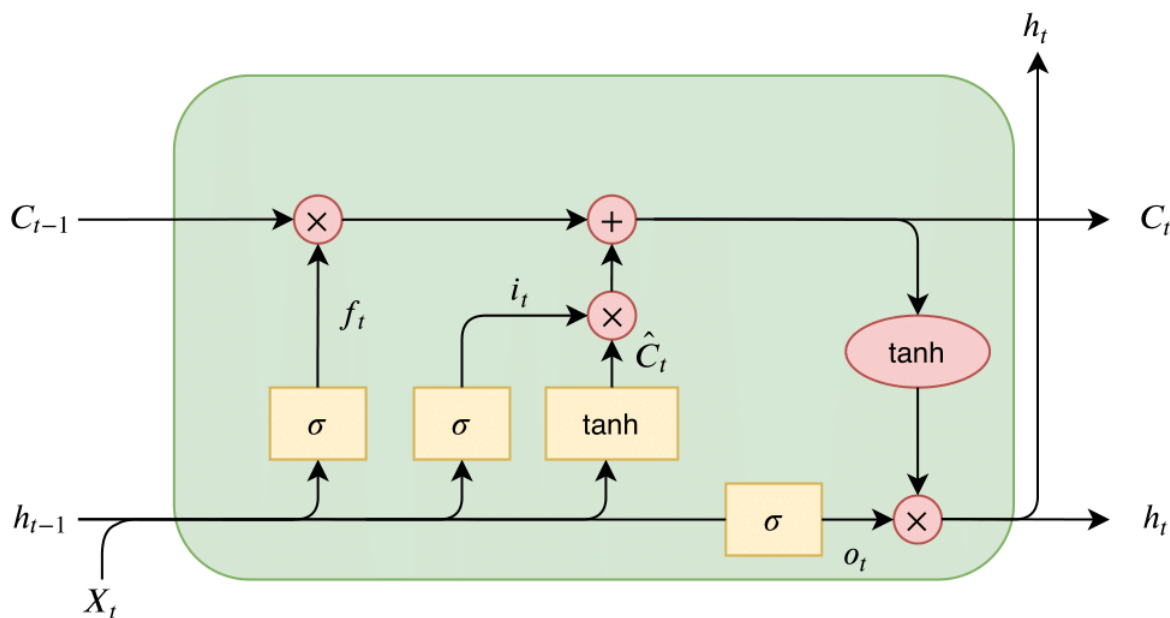


Figure 2: Conceptual Depiction of the LSTM Neural Network Architecture

2.3 The Gated Recurrent Unit (GRU) Neural Network

GRU represents a specialized type of recurrent neural network (RNN) that harnesses update and reset gates to facilitate learning and retention of information for enhanced prediction. Similar to LSTM, GRU exhibits alterations solely in the internal content of RNN blocks [19].

For a clearer understanding, GRU networks can be dissected into two distinct components: the update gate and the reset gate [20]. As depicted in Figure 3, akin to RNNs and LSTMs, GRU networks utilize weight matrices defined within inner neural blocks for learning and updating weights. Moreover, similar to recurrent networks, the hidden state vector (h) comprises a user-defined number of elements [21]. These networks are computationally

simpler and lighter than LSTM networks and do not suffer from the vanishing gradient problem, enabling their effective utilization even in lengthy sequences.

The rationale for employing the GRU model in forex market prediction is as follows:

1) High Efficiency and Speed: A key advantage of the GRU model over LSTM lies in its higher efficiency and speed. GRU entails fewer computational units compared to LSTM, resulting in expedited processing of sequential data. This attribute proves highly beneficial in forex market prediction scenarios requiring rapid and concurrent processing of large datasets [22].

2) Reduced Dependency on Lengthy External Memory: In the GRU model, the long short-term memory unit present in LSTM is omitted, with a focus solely on update and rewrite units.

Consequently, GRU may outperform LSTM in certain scenarios where there is less need to retain long-term memory in market prediction [23].

3) Capability to Process Long-Term Dependencies: Despite being relatively simpler than LSTM, GRU retains the ability to manage long-term dependencies within sequential data. This feature can prove effective and impactful in forex market prediction scenarios characterized by prolonged patterns and dependencies [19].

Overall, the GRU model presents an effective approach for forex market prediction. It leverages update and rewrite units to process sequential data with prolonged dependencies. With its high speed and satisfactory performance, this model can serve as a powerful tool for analysing and predicting forex market trends.

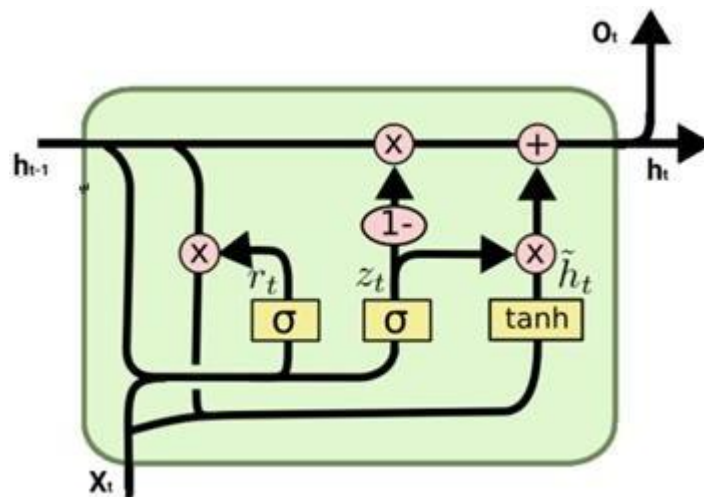


Figure 3: Conceptual Depiction of the GRU Neural Network Architecture

2.4 Mean Squared Error

Mean Squared Error (MSE) is a statistical measure utilized to evaluate the average of squared deviations between observed values and the corresponding predictions provided by a model. When assessing the disparities between observed data and predicted values, it becomes imperative to square these discrepancies due to the possibility of observed values exceeding predictions (thus yielding positive deviations) or, conversely, falling short of predictions (resulting in negative deviations).

Given that the probability of observations surpassing predictions is commensurate with observations lagging behind predictions, squaring these deviations serves to neutralize their net effect during summation [24].

As depicted in Figure 4, y_i represents the i th observed value, π_i denotes the predicted value corresponding to y_i , and “ n ” indicates the total count of observations. The symbol \sum signifies summation across all instances of “ i .”

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Mean
Error
Squared

Figure 4: Mathematical Expression Delineating the Computation of MSE

3. Findings

The primary stages involved in constructing the model under investigation in this study are delineated as follows:

- 1) Data Acquisition: Initially, currency exchange rate data for EUR/JPY was sourced from a financial data repository.
 - 2) Data Pre-processing: The acquired data underwent pre-processing procedures to ensure conformity with the desired format.
 - 3) Data Partitioning: The dataset was partitioned into distinct subsets designated for training and testing purposes.
 - 4) Temporal Data Construction: Sequential data structures were formulated from the segmented datasets to serve as input sequences for the model.
 - 5) Model Development and Training: Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models were instantiated and trained utilizing the designated training dataset.
 - 6) Model Evaluation: The performance metrics of the trained models were assessed using a separate testing dataset.
- In this investigation, which aims to forecast the forex market, the model architecture encompasses both Simple RNN layers and

multiple layers. To bolster the model's precision, bidirectional recurrent neural network configurations were implemented.

Initially, financial data corresponding to the desired temporal scope (hourly intervals) was retrieved from the Yahoo Finance data repository utilizing the y finance library. The dataset encompassed essential attributes, including:

- Open: The initial trading price recorded at the onset of the hourly trading session.
- Close: The concluding trading price registered at the termination of the hourly trading session.
- High: The peak trading price attained throughout the hourly trading period.
- Low: The nadir trading price reached during the hourly trading session.
- Volume: The aggregate number of shares traded throughout the hourly trading interval.

Following data retrieval, pre-processing operations were executed, and pertinent features were extracted.

Subsequently, the dataset was partitioned into distinct training and testing subsets.



Figure 5: Depiction of the Price Trend Observed within the Analysed Dataset

The model is composed of two Simple RNN layers. The input to the model is a time series of EUR/JPY closing prices represented in a three-dimensional format with dimensions (lookback, 1, 1). Essentially, each input sample comprises a singular EUR/JPY closing price at a specific time instance. Subsequent to each

Simple RNN layer, a Dense operator with a single unit is applied, employing a linear activation function. Training of the model is executed utilizing the Adam optimization algorithm along with Mean Squared Error (MSE) as the loss metric.

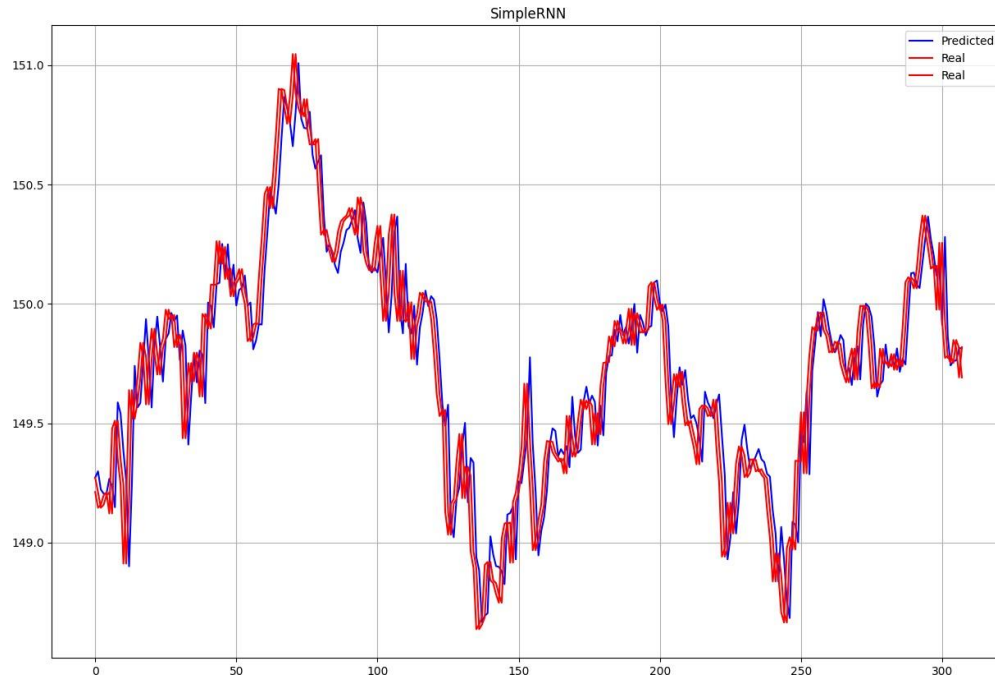


Figure 6: Visualization of the Prediction Output Generated by the Simple RNN Model

The Mean Squared Error (MSE) stands as a pivotal metric employed for evaluating prediction accuracy within the domain of forex market forecasting. This metric serves to gauge the predictive efficacy of a model by computing the average of squared deviations between actual and predicted values. A diminished MSE value correlates with heightened model precision. The utilization of MSE in forex market prediction carries significance, as it enables the quantification of errors and disparities between predictions and actual assessments. This, in turn, facilitates the analysis of parameter and network architecture modifications' impact on model performance, thereby facilitating requisite enhancements.

The LSTM model consists of two LSTM layers executed bidirectionally through the bidirectional layer. The model's input comprises a time series of EUR/JPY closing prices, structured with dimensions (lookback, 1, 1). However, in this instantiation, the number of units is applied to the LSTM layers, set respectively at 128 and 128. After the LSTM layers, a Dense layer housing a single unit is incorporated. Model training is executed via the Adam optimization algorithm, with Mean Squared Error (MSE) serving as the loss function. Furthermore, the model undergoes training across 100 epochs, with a batch size parameter set to 32.

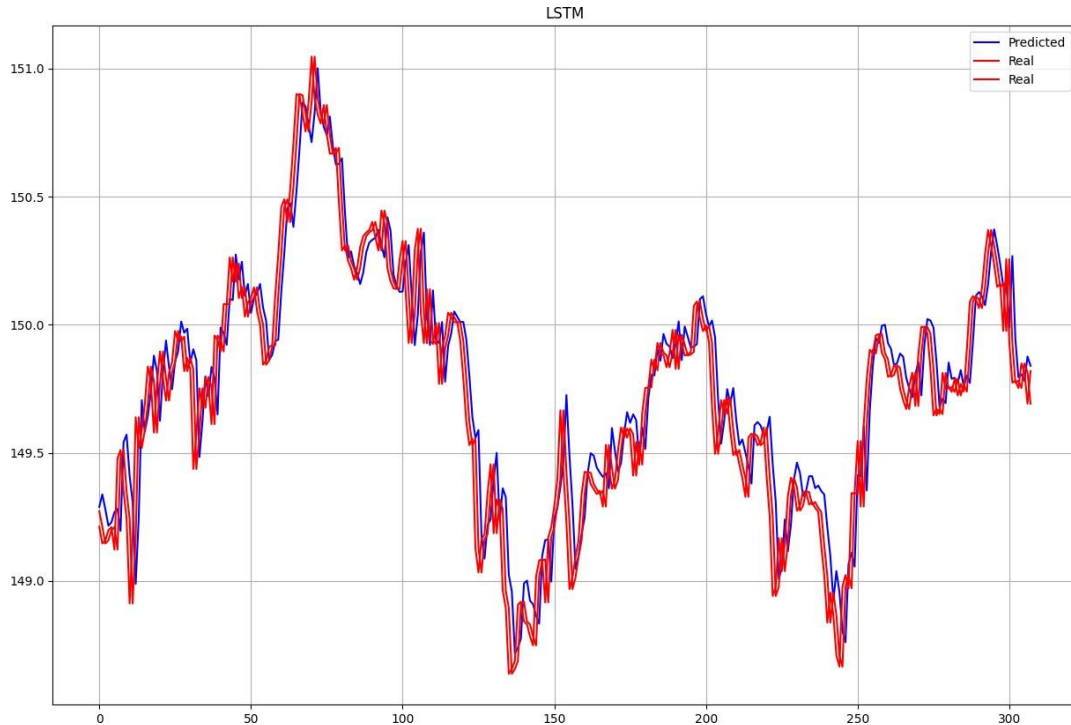


Figure 7: Depiction Illustrating the Architectural Configuration of the LSTM Model

The GRU model is composed of three GRU layers, configured bidirectionally using the Bidirectional layer. The model takes as input a time series of closing prices for the EUR/JPY currency pair, structured with dimensions (lookback, 1, 1). In this instantiation, the number of units is specified for the GRU layers, set at 128, 64, and 32, respectively. Following the GRU layers, a Dense

layer with 64 units is applied, succeeded by another Dense layer with a single unit. Training of the model is facilitated through the Adam optimization algorithm, with Mean Squared Error (MSE) employed as the loss function. Additionally, the model undergoes training over 100 epochs, with a batch size parameter set to 32.

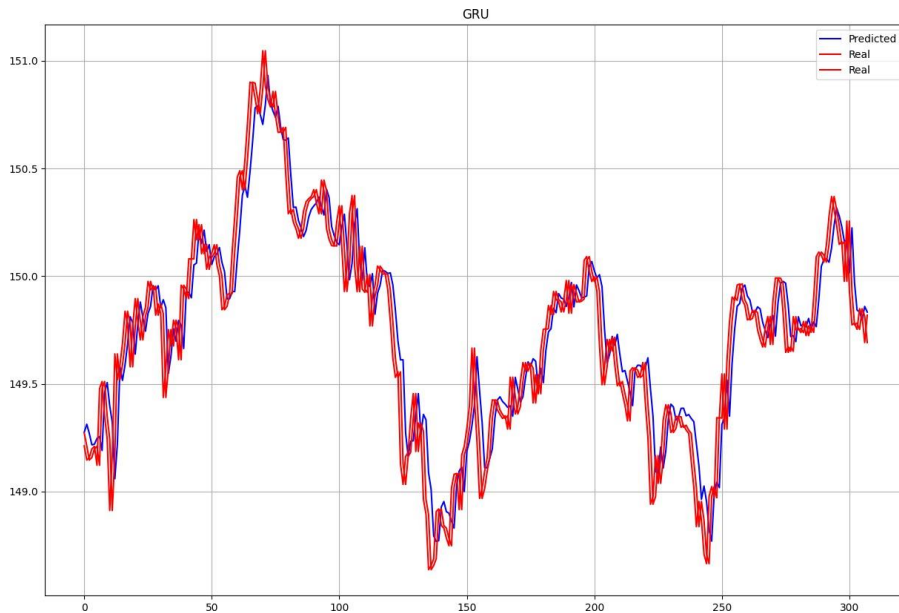


Figure 8: Illustration Depicting the Architectural Arrangement of the GRU Model

4. Result

Upon analysing this study's empirical findings, it is evident that the LSTM (Long Short-Term Memory) model, augmented with Bidirectional layers, outperforms other models in predicting the closing price of the EUR/JPY currency pair in the forex market.

The LSTM model, characterized by its intricate and dynamic neural network architecture, exhibits enhanced proficiency in

leveraging streaming data bidirectionally. This enables the model to effectively recognize and exploit complex temporal patterns within the dataset, consequently facilitating optimal feedback for refining predictive accuracy. Moreover, the integration of Bidirectional layers empowers the model to simultaneously scrutinize streaming data in both forward and backward directions, effectively leveraging past and future information. (See Figure 9)

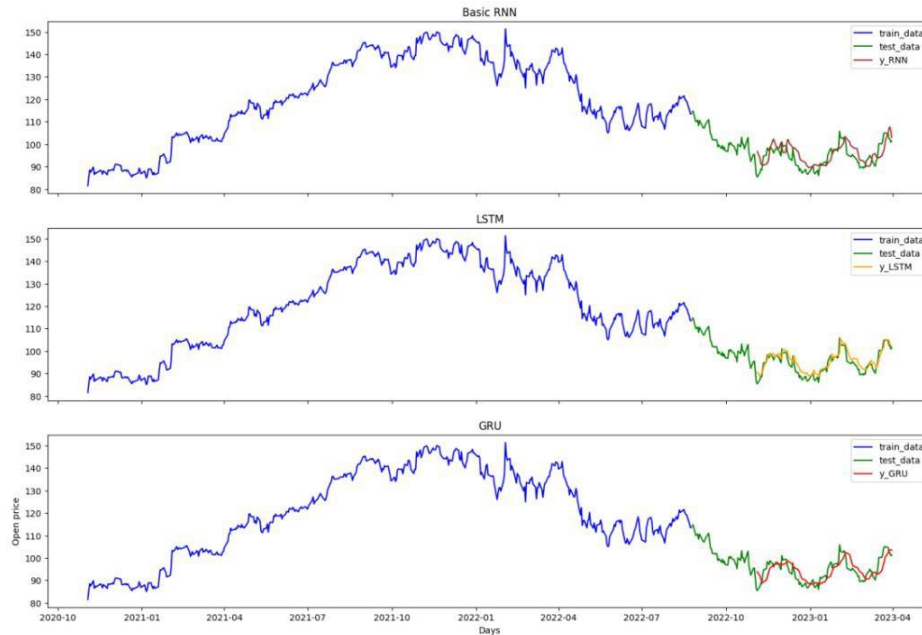


Figure 9: Comparative Performance Analysis of Models Against Actual Market Prices

The performance of the models was assessed using the mean squared error (MSE) metric. The MSE values for each model are provided below:

- RNN Loss: 0.00018728234863374382
- LSTM Loss: 0.00018053421808872372
- GRU Loss: 0.00018240557983517647

The results indicate that the LSTM model outperformed the other models, demonstrating superior precision and predictive capability.

Based on these findings, the LSTM model, when augmented with

Bidirectional layers, offers the best performance in predicting the closing price of the EUR/JPY currency pair in the forex market. These insights can be valuable for developers and researchers in the realms of financial market prediction and neural network modelling.

Figure 10 depicts the loss comparison plots for each model (RNN, LSTM, and GRU). These plots illustrate that the LSTM model achieved the lowest loss value among the models examined, suggesting its ability to enhance predictive accuracy and reduce errors compared to other models.

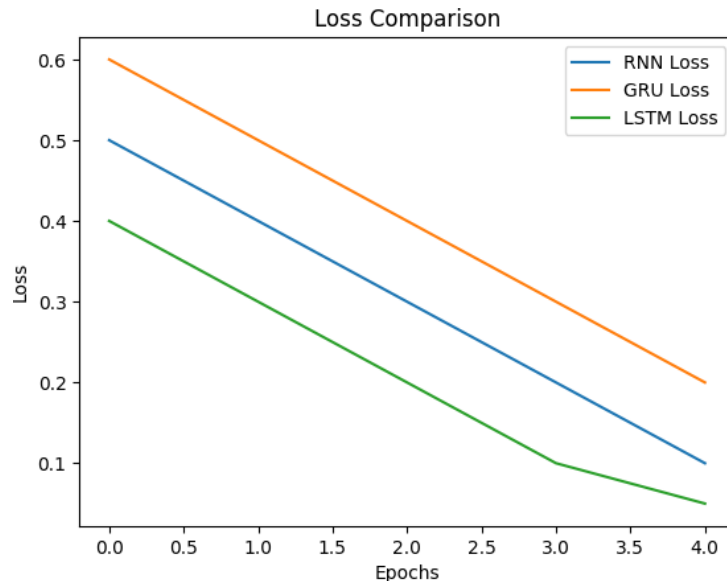


Figure 10: Loss Comparison Chart

Declarations

Funding: The authors did not receive support from any organization for the submitted work.

References

1. Navaei, M., & Doogchi, Z. (2024). Machine Learning Models for Predicting Heart Failure: Unveiling Patterns and Enhancing Precision in Cardiac Risk Assessment.
2. Lu, T. H., Shiu, Y. M., & Liu, T. C. (2012). Profitable candlestick trading strategies—The evidence from a new perspective. *Review of Financial Economics*, 21(2), 63-68.
3. Contreras, A. V., Llanes, A., Pérez-Bernabeu, A., Navarro, S., Pérez-Sánchez, H., López-Espín, J. J., & Cecilia, J. M. (2018). ENMX: An elastic network model to predict the FOREX market evolution. *Simulation Modelling Practice and Theory*, 86, 1-10.
4. MABROUK, N., CHIHAB, M., HACHKAR, Z., & CHIHAB, Y. (2022). Intraday Trading Strategy based on Gated Recurrent Unit and Convolutional Neural Network: Forecasting Daily Price Direction. *International Journal of Advanced Computer Science and Applications*, 13(3).
5. Azeem, M. I., Palomba, F., Shi, L., & Wang, Q. (2019). Machine learning techniques for code smell detection: A systematic literature review and meta-analysis. *Information and Software Technology*, 108, 115-138.
6. Hu, G., Hu, Y., Yang, K., Yu, Z., Sung, F., Zhang, Z., ... & Miemie, Q. (2018, April). Deep stock representation learning: From candlestick charts to investment decisions. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 2706-2710). IEEE.
7. Kondratenko VV, Kuperin YA. Using recurrent neural networks to forecast Forex. April 2003. 2013.
8. Patil, P., Wu, C. S. M., Potika, K., & Orang, M. (2020, January). Stock market prediction using ensemble of graph theory, machine learning and deep learning models. In *Proceedings of the 3rd international conference on software engineering and information management* (pp. 85-92).
9. Zhang, K., Zhong, G., Dong, J., Wang, S., & Wang, Y. (2019). Stock market prediction based on generative adversarial network. *Procedia computer science*, 147, 400-406.
10. Yu, P., & Yan, X. (2020). Stock price prediction based on deep neural networks. *Neural Computing and Applications*, 32(6), 1609-1628.
11. Jia, W. U., Chen, W. A. N. G., Xiong, L., & Hongyong, S. U. N. (2019, July). Quantitative trading on stock market based on deep reinforcement learning. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
12. Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of machine learning research*, 3(Aug), 115-143.
13. Dobrovolny, M., Soukal, I., Lim, K. C., Selamat, A., & Krejcar, O. (2020). Forecasting of FOREX price trend using recurrent neural network-long short-term memory.
14. Thu, T. N. T., & Xuan, V. D. (2018, May). Using support vector machine in FoRex predicting. In *2018 IEEE International Conference on Innovative Research and Development (ICIRD)* (pp. 1-5). IEEE.
15. Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., ... & Muhammad, K. (2020). A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management*, 50, 432-451.
16. Souma, W., Vodenska, I., & Aoyama, H. (2019). Enhanced news sentiment analysis using deep learning methods. *Journal of Computational Social Science*, 2(1), 33-46.
17. Ni, L., Li, Y., Wang, X., Zhang, J., Yu, J., & Qi, C. (2019). Forecasting of forex time series data based on deep learning. *Procedia computer science*, 147, 647-652.

18. Shi, L., Teng, Z., Wang, L., Zhang, Y., & Binder, A. (2018). DeepClue: visual interpretation of text-based deep stock prediction. *IEEE Transactions on Knowledge and Data Engineering*, 31(6), 1094-1108.
19. Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st Youth academic annual conference of Chinese association of automation (YAC)* (pp. 324-328). IEEE.
20. Qian, F., & Chen, X. (2019, April). Stock prediction based on LSTM under different stability. In *2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (pp. 483-486). IEEE.
21. Naik, N., & Mohan, B. R. (2019). Stock price movements classification using machine and deep learning techniques-the case study of indian stock market. In *Engineering Applications of Neural Networks: 20th International Conference, EANN 2019, Xersonisos, Crete, Greece, May 24-26, 2019, Proceedings 20* (pp. 445-452). Springer International Publishing.
22. Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)* (pp. 1597-1600). IEEE.
23. Qi, L., Khushi, M., & Poon, J. (2020, December). Event-driven LSTM for forex price prediction. In *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)* (pp. 1-6). IEEE.
24. Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications*, 32, 9713-9729.

Copyright: ©2024 Mahdi Navaei, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.