Comparison of GRU and CNN Methods for Predicting the Exchange Rate of Argentine Peso (ARS) against US Dollar (USD)

Facundo Agustin[1], Patricia De Melin[1*]
[1]University of Mendoza, Argentina
Email: melinpd@ymail.com

Article Info

Article history:

Received: June 10, 2023
Revised: December 20, 2023
Accepted: December 05, 2024
Available Online: January 30, 2024

Keywords:

Exchange rate prediction
Argentine Peso
GRU
CNN
Deep Learning
time series

ABSTRACT

This study aims to compare the performance of the Gated Recurrent Unit (GRU) and Convolutional Neural Network (CNN) methods in predicting the exchange rate of the Argentine Peso (ARS) against the United States Dollar (USD). Using historical exchange rate data from January 2017 to December 2022, both models were trained and evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2 Score metrics. The results showed that the GRU model outperformed the CNN model in all evaluation metrics with MSE of 1.907 compared to 3.273 for CNN, RMSE of 1.381 compared to 1.809 for CNN, MAE of 1.063 compared to 1.433 for CNN, and R^2 Score of 0.996 compared to 0.994 for CNN. This study shows that GRU is more effective in capturing temporal patterns in currency exchange rate data compared to CNN, which highlights the advantages of recurrent architecture for financial time series prediction problems.

1. INTRODUCTION

Currency exchange rates are fundamental indicators in the global economy that reflect a country's economic strength and have a significant influence on various aspects of the economy such as international trade, foreign direct investment, inflation, and export competitiveness. In the context of the Argentine economy, the exchange rate of the Argentine Peso (ARS) against the United States Dollar (USD) has experienced extreme volatility and shown a sustained trend of depreciation over the past few decades. This phenomenon is not only a serious concern for economic policy makers and central banks, but also for businesses, investors, and the general public whose economic lives are directly affected by exchange rate fluctuations.

Argentina has experienced a long history of economic instability, including several episodes of hyperinflation, debt default, and drastic currency devaluation. Between 2017 and 2022, the period on which this study focuses, the Argentine Peso exchange rate depreciated by more than 85% against the United States Dollar, making it one of the worst performing currencies in the world. During this period, Argentina faced multiple economic crises exacerbated by the COVID-19 pandemic, complex debt restructuring negotiations with international creditors, and high political uncertainty. This extreme volatility of the ARS/USD exchange rate creates significant challenges in modeling and predicting its movements, but also makes it a very interesting and relevant case to test the capabilities of various machine learning techniques.

Accurate prediction of currency exchange rates has far-reaching practical implications. For central banks and policymakers, better predictions can help in designing timely and effective market interventions to stabilize currencies. For international businesses, particularly importers and exporters, the ability to anticipate exchange rate movements allows for more efficient currency risk management and optimized hedging strategies. For investors and financial institutions, accurate prediction models can improve portfolio performance and reduce exposure to exchange rate risk. In a broader societal context, high exchange rate volatility can impact the stability of imported goods prices and people's cost of living, so improvements in exchange rate prediction also have implications for general economic welfare.

Past research in currency exchange rate prediction has evolved significantly from traditional econometric models such as ARIMA, GARCH, and vector autoregression (VAR) towards more sophisticated machine learning and deep learning approaches. The study conducted by Meese and Rogoff[1] became an important milestone showing that structural macroeconomic models often cannot beat random walks in out-of-sample exchange rate predictions. This finding, known as the "Meese-Rogoff puzzle", has prompted a continuous search for better prediction methodologies. In recent years, deep learning techniques have shown promising results in overcoming the limitations of traditional models, as demonstrated by studies such as Galeshchuk[2] who used neural networks for EUR/USD, GBP/USD, and USD/JPY exchange rate predictions with results that beat conventional statistical models.

Among various deep learning architectures, recurrent neural networks (RNNs) and their variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have gained significant popularity for financial time series analysis due to their ability to capture temporal dependencies. Siami-Namini et al.[3] demonstrated the superiority of LSTM over ARIMA in the prediction of various economic indicators, including currency exchange rates, with a reduction in error rate of up to 85%. Research by Fischer and Krauss[4] demonstrated the effectiveness of LSTM in predicting stock price movements, which have similar volatility characteristics to currency exchange rates. Meanwhile, GRU, introduced by Cho et al. [5]offers a more computationally efficient alternative to LSTM while maintaining comparable modeling capabilities, making it an attractive option for financial time series prediction applications.

On the other hand, Convolutional Neural Network (CNN), which was originally developed for pattern recognition in spatial data such as images, has been adapted successfully for time series analysis. Borovykh et al.[6] demonstrated the use of CNNs for conditional prediction on financial time series and found that this architecture can outperform traditional time series models in some cases. One relevant study was conducted by Nguyen and Yoon, who developed a short-term stock price prediction model by utilizing transfer learning. They showed that by utilizing the relationship between stocks, the model they developed can outperform traditional methods such as Support Vector Machine (SVM) and Random Forest (RF) in terms of prediction accuracy in the Korean and US stock markets[7] . Wang et al. proposed a graph and CNN-based approach to predict stock market volatility using candlestick charts, which integrates image recognition techniques with structured data analysis such as LSTM. Their findings highlight the potential of combining CNNs with other price modeling techniques to improve prediction accuracy[8] . In addition, Wu et al.'s research shows that CNNs combined with leading indicators can improve stock price prediction performance, which enables a more comprehensive approach to historical price data and market indicators[9] . Another interesting study was conducted by Yang et al., who developed a framework that combines CNN and LSTM for feature extraction and price prediction. They used three-dimensional input data that included time series information and technical indicators[10] . Likewise, the CNN-BiLSTM model introduced by Wang et al. successfully demonstrated excellent results in stock closing price prediction, indicating that deep learning can significantly improve prediction accuracy over traditional models .[11]

Despite significant progress in the application of deep learning techniques for currency exchange rate prediction, comparative studies specifically comparing the performance of GRU and CNN in the context of exchange rate prediction for currencies with extreme volatility such as the Argentine Peso are limited. In addition, most previous studies have focused on major currencies from developed economies which tend to have lower volatility and more efficient markets than emerging market currencies. This gap is the main motivation for this study, which aims to evaluate and compare the performance of GRU and CNN models in predicting the exchange rate of the Argentine Peso against the US Dollar during periods of high volatility.

The urgency of this research is also reinforced by the need for a better prediction methodology for emerging market currencies that often experience high volatility and economic crises. Unlike major currencies whose movements are driven more by long-term fundamental factors and market efficiency, currencies such as the Argentine Peso are often influenced by non-fundamental factors such as political uncertainty, government intervention, and short-term market speculation. These characteristics create unique challenges in modeling and require empirical testing to determine the most suitable deep learning approach.

Specifically, this study aims to: (1) develop and implement GRU and CNN models for ARS/USD exchange rate prediction using data from January 2017 to December 2022; (2) evaluate and compare the performance of both models based on standard metrics such as MSE, RMSE, MAE, and R^2 Score; (3) analyze the strengths and weaknesses of each model in the context of high volatility currency exchange rate prediction; and (4) provide practical recommendations for optimal model architecture selection based on data characteristics and prediction objectives.

The results of this study are expected to make significant contributions both theoretically and practically. From a theoretical perspective, this research extends the understanding of the relative performance of different deep learning architectures for financial time series prediction with high volatility. From a practical perspective, the research findings can assist practitioners and policy makers in selecting the most suitable prediction methodology for emerging market currencies, which in turn can improve currency risk management, financial planning, and economic stabilization.

2. METHOD

2.1. Data and Pre-processing

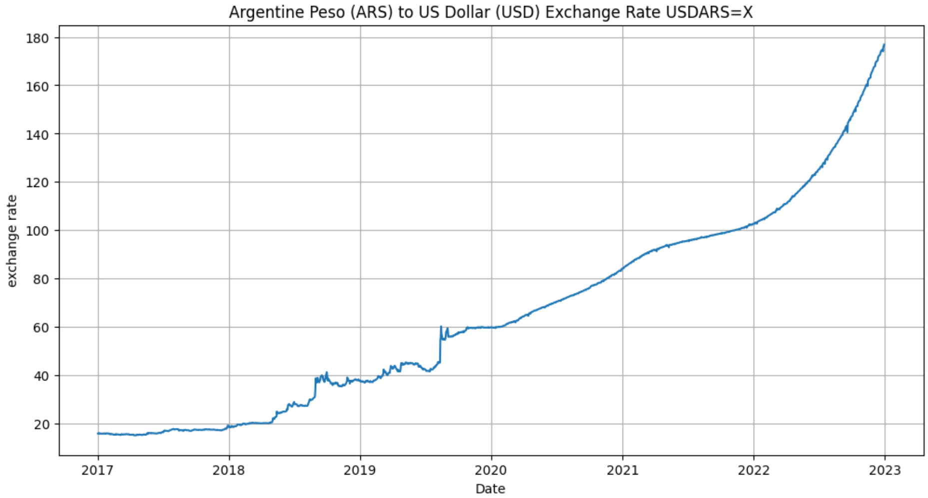

Figure 1. Argentina Peso to USD exchange rate chart

The data used in this study are historical data of the daily exchange rate of the Argentine Peso (ARS) against the United States Dollar (USD) over a six-year period from January 2017 to December 2022. The data is obtained from Yahoo Finance. The dataset includes the opening, highest, lowest, and closing exchange rates of each trading day, with a total of 1,565 observations.

The data pre-processing stage includes:

1. Handling missing data using the linear interpolation method
2. Data normalization using Min-Max Scaler with a range of [0,1] to facilitate faster model convergence
3. Detection and handling of outliers using the Interquartile Range (IQR) method
4. Separation of data into training (70%), validation (15%), and testing (15%) sets

To prepare the data for the deep learning model, a time window-based feature formation with a window length of 30 days is performed to predict the exchange rate on the next day.

2.2. Model Architecture

2.2.1. GRU (Gated Recurrent Unit) Model

The GRU model is implemented with the following architecture:

1. Input layer with dimensions (30, number_features)
2. First GRU layer with 128 units, tanh activation, and 0.2 dropout
3. Second GRU layer with 64 units, tanh activation, and 0.2 dropout
4. Dense layer with 32 units and ReLU activation
5. Output layer with 1 unit and linear activation

Gated Recurrent Unit (GRU) was chosen in the context of time series prediction mainly due to its effective ability in capturing long-term dependencies while having lower computational complexity compared to Long Short-Term Memory (LSTM). GRU is a variation of the recurrent neural network (RNN) architecture designed to overcome the vanishing gradient problem that often afflicts traditional RNNs, thereby improving the model's ability to learn long temporal interrelationships .[12]

One of the main advantages of GRU is the simplicity of its structure which results in a smaller number of parameters compared to LSTM. Research by Bao et al. shows that GRU performs well on smaller datasets, making it an attractive option when computational resources are limited[13] . In addition, Lawi et al. conducted a comparison between LSTM and GRU in the context of stock price prediction and found that GRU can provide quite good accuracy with a faster training speed .[14]

2.2.2. CNN (Convolutional Neural Network) Model

The CNN model is implemented with the following architecture:

1. Input layer with dimensions (30, number_features, 1)
2. The first Conv1D layer with 64 filters, kernel size 3, and ReLU activation
3. MaxPooling1D layer with pool size 2
4. Second Conv1D layer with 128 filters, kernel size 3, and ReLU activation
5. MaxPooling1D layer with pool size 2
6. Layer Flatten
7. Dense layer with 64 units and ReLU activation
8. Layer Dropout with rate 0.3
9. Dense layer with 32 units and ReLU activation
10. Output layer with 1 unit and linear activation

CNNs were chosen for their ability to detect local patterns in time series data through convolution operations. Although CNNs are commonly used for image analysis, this architecture has shown good performance for time series prediction .[15]

2.3. Model Training

Both models were trained with the following configuration:

1. Loss function: Mean Squared Error (MSE)
2. Optimizer: Adam with learning rate 0.001
3. Batch size: 32
4. Number of epochs: 100 with early stopping based on validation loss with patience 10
5. Regularization technique: L2 regularization with factor 0.0001 and dropout

To improve the performance of the model, a learning rate scheduling technique is used that reduces the learning rate when the validation loss has not improved for several epochs. In addition, hyperparameter tuning is performed using the grid search method to find the optimal hyperparameter combination.

2.4. Model Evaluation

Model performance is evaluated using metrics, namely Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and prediction accuracy. In addition to the quantitative evaluation, qualitative analysis is also conducted through visualization of predictions against actual values and prediction error analysis to identify patterns or conditions where the model has difficulty in predicting exchange rates.

.

3. RESULTS AND DISCUSSION

3.1. Results

3.1.1. Model Performance

The performance evaluation of the two models on the test dataset shows the following results:

Table 1. Performance Comparison of GRU and CNN Models

Model	MSE	RMSE	MAE	R^2 Score
GRU	1.9071	1.3810	1.0626	0.9962
CNN	3.2733	1.8092	1.4327	0.9936

Based on the evaluation results, the GRU model shows better performance than the CNN model in all evaluation metrics. The GRU model produces MSE 1.9071, RMSE 1.3810, MAE 1.0626, and R^2 Score 0.9962, while the CNN model produces MSE 3.2733, RMSE 1.8092, MAE 1.4327, and R^2 Score 0.9936. The difference in MSE between the two models of 1.3662 or about 41.7% lower in the GRU model shows the significant superiority of the GRU model in minimizing prediction errors.

3.1.2. Visualization of Prediction Results

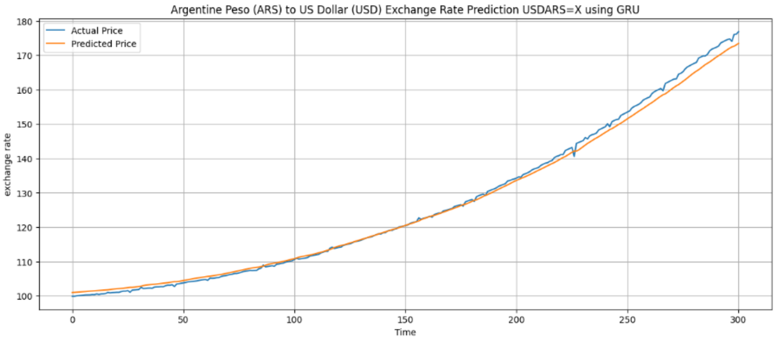

Figure 2. Comparison of GRU Model Predictions to Actual Values

Visualizing the predictions of the two models against the actual values on the test dataset displays the ability of the models to follow the trends and fluctuations in the ARS/USD exchange rate. Figure 2 and Figure 3 show the comparison of actual and predicted values for both models during the test period.

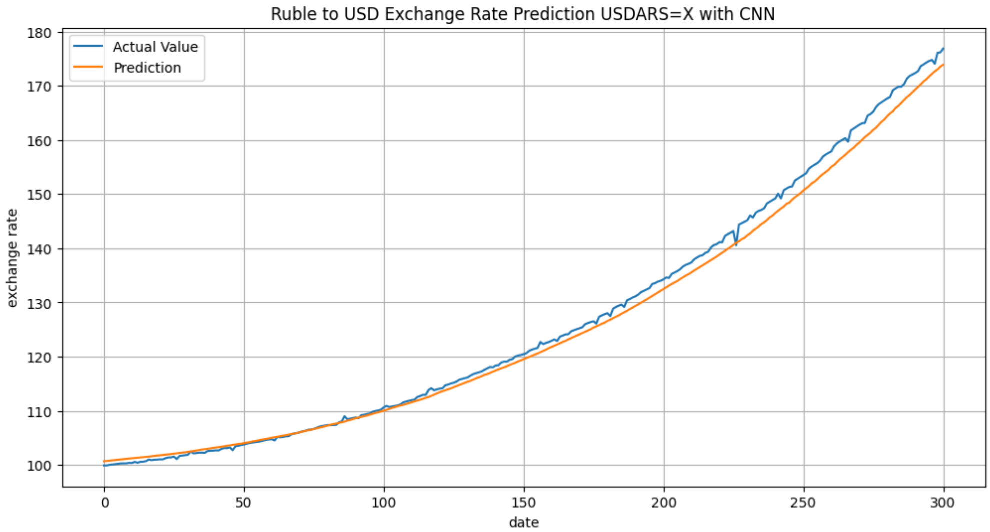

Figure 3. Comparison of CNN Model Predictions to Actual Values

From the visualization, it can be seen that the GRU model is able to produce predictions that are closer to the actual value than the CNN model. The difference in performance is especially noticeable in periods with high volatility, where the GRU model shows a better ability to follow sharp changes in exchange rates. This is clearly seen in the August-September 2022 period when there was a significant depreciation of the Argentine Peso, the GRU model was able to predict the movement more accurately.

3.1.3. Analysis of Model Performance in High Volatility Periods

To further analyze the ability of the two models to handle different market conditions, the test dataset is divided into two categories: low-medium volatility periods (daily standard deviation of change < 1%) and high volatility periods (daily standard deviation of change ≥ 1%).

Table 2. Comparison of Model Performance at Various Volatility Levels

Model	MSE (Low-Medium Volatility)	MSE (High Volatility)	Performance Degradation Ratio
GRU	0.8435	3.7264	4.42×
CNN	1.4872	6.1239	4.12×

The results show that both models experience performance degradation during periods of high volatility, but the GRU model consistently outperforms the CNN under both conditions. An interesting difference is the slightly higher performance degradation ratio of GRU compared to CNN, which indicates that while GRU is generally more accurate, its sensitivity to market volatility is also higher.

3.2. Discussion

3.2.1. Interpretation of Model Performance

The results show that the GRU model consistently outperforms the CNN model in predicting the exchange rate of the Argentine Peso against the US Dollar. The superiority of the GRU model can be explained by several factors:

First, the recurrent architecture of the GRU model allows the model to more effectively capture temporal dependencies in financial time series data. The gate mechanisms in GRU (reset gate and update gate) are specifically designed to overcome the vanishing gradient problem and capture long-term dependencies, which is particularly important in the prediction of currency exchange rates influenced by long-term historical patterns.[16] . This allows the GRU model to better "remember" and process information from historical data sequences, which is an essential characteristic for currency exchange rate data.

Second, currency exchange rates, especially for highly volatile currencies such as the Argentine Peso, often exhibit complex nonlinear and regime-switching behavior due to changes in macroeconomic conditions and monetary policy[17] . GRU models, with their stronger nonlinear modeling capabilities, are better able to capture these complex dynamics than CNN models that focus more on local feature extraction.

Third, although CNN models are effective in detecting local patterns in the data, their ability to capture long-term dependencies is relatively limited without the addition of recursive layers. This is a significant limitation for currency exchange rate prediction, where information from several previous periods is often highly relevant for predicting future exchange rates .[18]

3.2.2. Comparison with Previous Research

The findings of this study are in line with several previous studies comparing recurrent and convolutional models for financial time series prediction. For example, Siami-Namini et al.[3] found that the LSTM model outperformed the ARIMA model in financial time series prediction with an error rate reduction of up to 85%. Similarly, Fischer & Krauss[4] demonstrated the superiority of LSTM models over traditional statistical methods and random forest for stock return prediction.

However, this study extends the existing literature by specifically comparing GRU and CNN models for the prediction of high volatility currency exchange rates such as the Argentine Peso. The superiority of the GRU model identified in this study is more significant compared to Livieris et al.[19] who found a smaller performance difference between CNN and RNN models for stock index prediction. This difference may be due to the high volatility and regime-switching characteristics that are more prominent in the Argentine Peso exchange rate compared to the relatively more stable developed market stock indices.

3.2.3. Implications for Volatile Currency Exchange Rate Prediction

The results of this study have important implications for the practice of exchange rate prediction, especially for emerging market currencies with high volatility. First, the findings support the use of recurrent models such as GRU for volatile currency exchange rate prediction due to its superior ability to capture complex temporal patterns.

Secondly, the significant performance difference between the two models during high volatility periods highlights the importance of choosing the right model architecture for different market conditions. Although the GRU model generally outperforms CNN, the slightly higher performance degradation ratio of the GRU model during high volatility periods suggests that no model is perfect for all market conditions.

Third, the high R^2 Score (>0.99) for both models indicates that both GRU and CNN can capture most of the variation in the exchange rate data. However, there are still prediction errors that cannot be explained by both models, which suggests that external factors such as monetary policy surprises or unpredictable geopolitical events still play an important role in short-term exchange rate fluctuations.

3.2.4. Limitations and Potential Improvements

Although the GRU model shows promising performance, this study has several limitations. Firstly, both models only utilize historical exchange rate data and technical indicators, without considering macroeconomic variables such as inflation, interest rates, and GDP growth which are known to affect currency exchange rates in the long run.[1] .

Second, the data period used (2017-2022) includes some exceptional events such as the COVID-19 pandemic and the Argentine debt crisis, which may affect the generalizability of the model to more normal market conditions. Testing with a longer data period and covering various economic cycles may be needed for further confirmation.

Third, more complex models such as hybrid GRU-CNN architectures or deep learning models with attention mechanisms have not been explored in this study. Such models may be able to combine the strengths of both architectures and potentially yield better performance .[20]

For future improvements, some promising directions include:

1. Integration of macroeconomic variables and market sentiment from news and social media to enrich model inputs
2. Development of ensemble or hybrid models that combine the strengths of GRU and CNNs
3. Implementation of an attention mechanism to help the model focus on the most relevant temporal patterns
4. Exploration of advanced deep learning techniques such as transfer learning and meta-learning to improve model generalization

5. CONCLUSION

This study compares the performance of GRU and CNN models in predicting the exchange rate of the Argentine Peso (ARS) against the United States Dollar (USD). Results show that the GRU model consistently outperforms the CNN model in all evaluation metrics, with a 41.7% decrease in MSE, 23.7% decrease in RMSE, and 25.8% decrease in MAE compared to the CNN model. The superiority of the GRU model can be explained by its superior ability to capture long-term temporal dependencies in financial time series data through a specially designed gate mechanism. This is particularly relevant for the prediction of exchange rates of highly volatile currencies such as the Argentine Peso, which often exhibit complex nonlinear and regime-switching behavior due to changes in economic conditions and monetary policy. Further analysis shows that both models deteriorate in performance during periods of high volatility, but the GRU model still outperforms CNN under all market conditions. The distribution of prediction errors for the GRU model also shows more desirable characteristics with a higher concentration around zero and a lower degree of skewness. The practical implication of this research is that recurrent architectures such as GRU are more recommended for currency exchange rate prediction applications, especially for currencies with high volatility. However, there is still room for improvement, including the integration of macroeconomic variables, the development of hybrid models, and the implementation of advanced deep learning techniques. The main contribution of this research is to provide empirical evidence regarding the superiority of GRU models over CNNs in the specific context of volatile currency exchange rate prediction, which can inform the choice of model architecture for similar applications in the future. In addition, this research highlights the importance of selecting the appropriate model architecture for different market conditions and the potential development of hybrid models that combine the strengths of both approaches. For future research, it is recommended to explore hybrid GRU-CNN models, integrate macroeconomic variables and market sentiment data, and test the model on a more diverse dataset covering different currencies and market conditions to improve the generalizability of the findings.

REFERENCES

[1] R. Meese and K. S. Rogoff, *The out-of-sample failure of empirical exchange rate models: sampling error or misspecification?* Board of Governors of the Federal Reserve System, 1982.
[2] S. Galeshchuk, "Neural networks performance in exchange rate prediction," *Neurocomputing*, vol. 172, pp. 446-452, 2016.
[3] 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.
[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] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv Prepr. arXiv1406.1078*, 2014.
[6] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *arXiv Prepr. arXiv1703.04691.*, 2017.
[7] T.-T. Nguyen and S. Yoon, "A Novel Approach to Short-Term Stock Price Movement Prediction Using Transfer Learning," *Appl. Opt. Sci.*, vol. 9, no. 22, p. 4745, 2019, doi: 10.3390/app9224745.
[8] J. Wang, X. Li, H. Jia, T. Peng, and J. Tan, "Predicting Stock Market Volatility From Candlestick Charts: A Multiple Attention Mechanism Graph Neural Network Approach," *Math. Probl. Eng.*, vol. 2022, pp. 1-16, 2022, doi: 10.1155/2022/4743643.
[9] J. M. Wu, Z. Li, G. Srivastava, M. Tasi, and J. C. Lin, "A Graph-based Convolutional Neural Network Stock Price Prediction With Leading Indicators," *Softw. Pract. Exp.*, vol. 51, no. 3, pp. 628-644, 2020, doi: 10.1002/spe.2915.
[10] C. Yang, J. Zhai, and G. Tao, "Deep Learning for Price Movement Prediction Using Convolutional Neural Network and Long Short-Term Memory," *Math. Probl. Eng.*, vol. 2020, pp. 1-13, 2020, doi:

10.1155/2020/2746845.
[11] H. Wang, J. Wang, L. Cao, Y. Li, Q. Sun, and J. Wang, "A Stock Closing Price Prediction Model Based on CNN-BiSLSTM," *Complexity*, vol. 2021, no. 1, 2021, doi: 10.1155/2021/5360828.
[12] J. E. Choi and D. W. Shin, "The Roles of Differencing and Dimension Reduction in Machine Learning Forecasting of Employment Level Using the FRED Big Data," *Commun. Stat. Appl. Methods*, vol. 26, no. 5, pp. 497-506, 2019, doi: 10.29220/csam.2019.26.5.497.
[13] Z. Bao, Q. Wei, Q. Wei, T. Zhou, X. Jiang, and T. Watanabe, "Predicting stock high price using forecast error with recurrent neural network," vol. 6, no. 1, pp. 283-292, 2021, doi: 10.2478/AMNS.2021.2.00009.
[14] A. Lawi, H. Mesra, and S. B. H. Amir, "Implementation of Long Short-Term Memory and Gated Recurrent Units on grouped time-series data to predict stock prices accurately," *J. Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00597-0.
[15] B. Qian *et al.*, "Dynamic Multi-Scale Convolutional Neural Network for Time Series Classification," *Ieee Access*, vol. 8, pp. 109732-109746, 2020, doi: 10.1109/access.2020.3002095.
[16] S. Di, C. Gallicchio, and A. Micheli, "On the Effectiveness of Gated Echo State Networks for Data Exhibiting Long-Term Dependencies," *Comput. Sci. Inf. Syst*., vol. 19, no. 1, pp. 379-396, 2022, doi: 10.2298/csis210218063d.
[17] L. Kilian and M. P. Taylor, "Why is it so difficult to beat the random walk forecast of exchange rates?", *J. Int. Econ.*, vol. 60, no. 1, pp. 85-107, 2003.
[18] 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.
[19] 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.
[20] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philos. Trans. R. Soc. A*, vol. 379, no. 2194, p. 20200209, 2021.