International Journal of Artificial Intelligent and Informatics
Vol. 3, No. 1, January 2025, pp. 24~ 29
ISSN: 2622-626X, DOI: http://doi.org/10.33292/ijarlit.v3i1.43

Comparison of LSTM and GRU Methods for Predicting Gold Exchange Rate against US Dollar

Dušan Bohovic
[1]University of Novi Pazar, Serbia
Email: bohovicd@gmail.com

Article Info

Article history:

Received: June 15, 2024
Revised: August 10, 2024
Accepted: December 10, 2024
Available Online: January 30, 2025

Keywords:

gold exchange rate prediction
deep learning
LSTM
GRU
financial time series

ABSTRACT

This study aims to compare the performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models in predicting the gold exchange rate against the United States Dollar (USD). Using time series data from Yahoo Finance for the period 2017-2023, we evaluate and compare the two models based on comprehensive evaluation metrics. The results show that the GRU model performs better in several important metrics, especially in terms of Root Mean Square Error (RMSE) on the test data (26.41 compared to 27.54 on LSTM) and higher coefficient of determination (R^2) on the test data (0.9004 compared to 0.7825 on LSTM). These findings indicate that the GRU model has better generalization ability for gold to USD exchange rate prediction, although both models show very high accuracy rates above 98% on the test data.

1. INTRODUCTION

Gold has long been considered a stable investment instrument and a hedge against inflation and economic instability. The exchange rate of gold against the United States Dollar (USD) is an important indicator in global financial markets that influences various investment and monetary policy decisions. Gold's price fluctuations, which are influenced by various macroeconomic, geopolitical and market sentiment factors, make predicting its exchange rate a complex but crucial challenge for investors, commodity traders and policymakers.

In recent years, machine learning approaches have shown promising capabilities in predicting financial time series, including commodity prices such as gold. Among various machine learning techniques, recurrent neural network models (RNNs) have emerged as an effective approach for time series analysis due to their ability to capture temporal dependencies in data. In particular, the Long Short-Term Memory (LSTM) introduced by Hochreiter and Schmidhuber[1] and the Gated Recurrent Unit (GRU) proposed by Cho et al.[2] are two popular RNN architectures for time series modeling.

LSTM and GRU are both designed to overcome the vanishing gradient problem that often occurs in traditional RNNs when handling long-term dependencies. LSTM uses a complex system of gates including input gates, forget gates, and output gates to control the flow of information, while GRU uses a simpler design with only two gates - a reset gate and an update gate. Although GRU has a simpler structure, some research shows that its performance can be equivalent or even better than LSTM in certain contexts .[3], [4]

The use of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) in time series forecasting has shown significant progress thanks to their ability to handle the vanishing gradient problem that often occurs in traditional Recurrent Neural Network (RNN) models. LSTMs, with their complex gate structure, allow these models to store long-term information and recognize patterns in time series effectively[5], [6] . A study shows that LSTM and GRU are able to reduce type II and type I errors better than the traditional RNN, confirming their ability for more precise forecasting .[5]

GRU, as a variant of RNN, offers a simpler structure by combining input gates and forget gates into update gates, which reduces the computational burden while still maintaining effective performance in time series forecasting[7] . In the context of forecasting, research shows that GRU can automatically retain useful information and leave irrelevant information in dynamic sequence data[8] . Other research also highlights that GRU has excellent performance, almost on par with LSTM, in a variety of time series settings, including the prediction of reservoir parameters using well logging data .[9]

The implementation of GRU and LSTM in different domains, such as short-term load monitoring and air quality prediction, shows their great potential in various practical applications. For example, GRU models used for short-term load forecasting show better results than other simpler models, thanks to their ability to handle multi-source data[10], [11] . In addition, GRU also shows outstanding performance in the prediction of optimal time to maintain temperature conditions in energy-efficient buildings through more complex learning models[12] . On the other hand, LSTM has proven to excel in stock price prediction, where a study showed that its prediction accuracy improved substantially thanks to its strength in learning from long and complex time series data .[13], [14]

Both models have gained wide attention in research and practical applications, illustrating the transition from traditional forecasting techniques such as ARIMA towards more efficient deep learning-based approaches[15] . In summary, both LSTM and GRU are at an important position in the evolution of time series forecasting models, enabling better accuracy and wider applications in various fields, such as economics, health, and environment .[16], [17]

Although a number of studies have compared the effectiveness of LSTM and GRU in various applications, such as natural language processing[18] , sentiment analysis[19] , and weather forecasting[20] , there have not been many comprehensive studies that specifically compare the performance of these two models in the context of gold to USD exchange rate prediction using recent data covering periods of high volatility such as the COVID-19 pandemic and fluctuating global economic conditions.

This study aims to fill the gap by conducting a thorough comparative evaluation of the LSTM and GRU models in predicting the gold to USD exchange rate using data from Yahoo Finance for the period 2017-2023. We evaluate the performance of both models based on a comprehensive range of evaluation metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and prediction accuracy.

The results of this study are expected to provide valuable insights for practitioners and researchers in the field of computational finance on more suitable models for gold exchange rate prediction. In addition, the findings may assist investors and market analysts in developing more effective trading and investment strategies based on more accurate predictions.

2. METHOD

2.1. Data

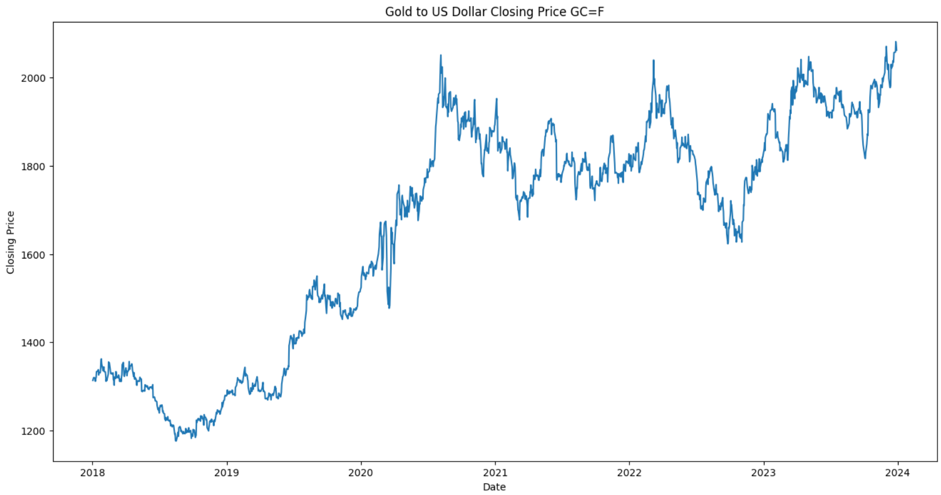

Figure 1. Gold value movement against USD

In this study, we used time series data of the gold exchange rate against the United States Dollar (USD) obtained from Yahoo Finance for the period 2017 to 2023. This data includes the daily closing price of gold in USD per troy ounce. The choice of this time period allows the model to study patterns across a wide range of market conditions, including periods of high volatility such as the COVID-19 pandemic and fluctuating global economic conditions.

The dataset was divided into two parts: training data and testing data in a ratio of 80:20. A total of 80% of the initial data is used to train the model, while the last 20% of the data is used to evaluate the performance of the model. This division allows for a comprehensive evaluation of the generalization ability of the model.

2.2. Data

Before being used for model training, the data goes through several pre-processing stages:

1. Data Cleaning: We check and remove missing values and outliers in the dataset.
2. Normalization: Data is normalized using the Min-Max Scaler method to adjust all features to the range [0, 1], which helps prevent dominance of certain features and accelerates convergence during training.
3. Sequence Formation: The data is converted into a sequence-to-sequence format, where each sequence consists of the previous 60 days of observations (features) and the gold exchange rate on the following day (target).

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. First LSTM layer with 100 units, using tanh activation and sigmoid recurrence
2. Dropout (0.2) to prevent overfitting
3. Second LSTM layer with 50 units
4. Dropout (0.2)
5. Dense layer with 1 unit to generate gold exchange rate prediction

2.2.1. GRU Model

The GRU model used has a similar structure to the LSTM model:

1. First GRU layer with 100 units, using tanh activation and sigmoid recurrence
2. Dropout (0.2)
3. Second GRU layer with 50 units
4. Dropout (0.2)
5. Dense layer with 1 unit to generate prediction

2.3. Model Training

Both models were trained with the following configuration:

1. Optimizer: Adam with learning rate 0.001
2. Loss function: Mean Squared Error (MSE)
3. Batch size: 32
4. Epochs: 100 with Early Stopping (patience=10) to avoid overfitting
5. Validation split: 20% of training data

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. Results

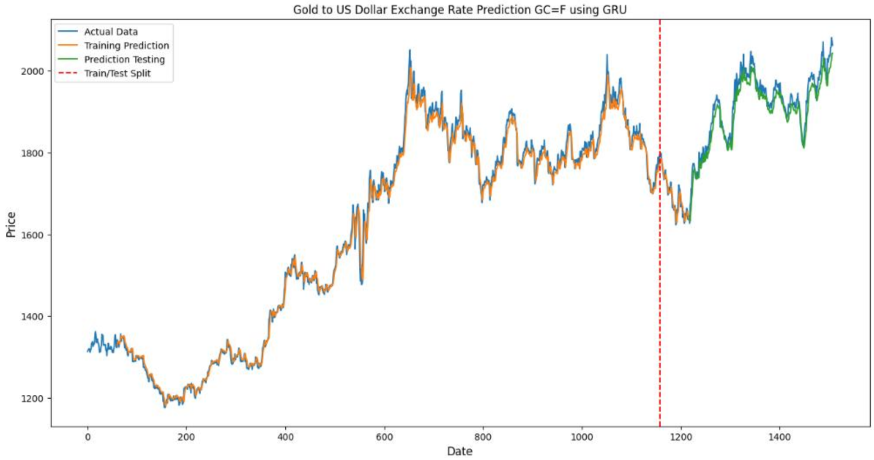

Figure 2. Comparison between actual data and GRU Model prediction

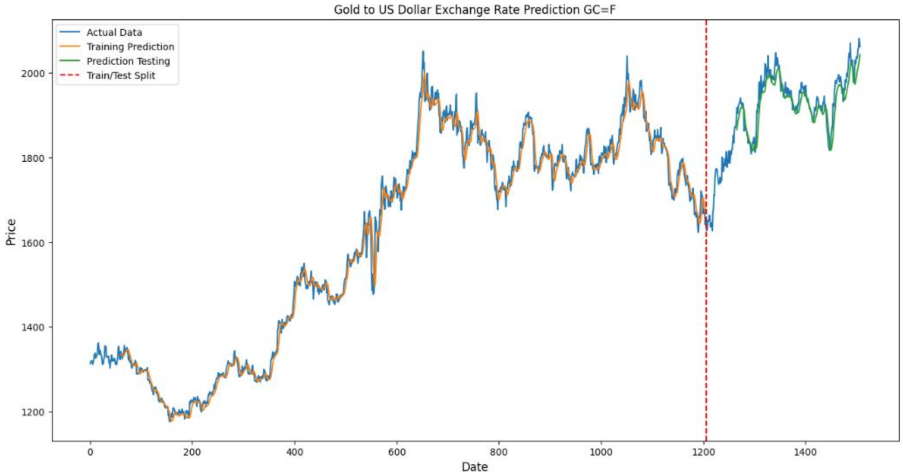

Figure 3. Comparison between actual data and LSTM Model prediction

The evaluation results of the LSTM and GRU models on training and testing data are presented in Table 1.

Table 1. Comparison of LSTM and GRU Model Evaluation Metrics

Metrics	LSTM (Training)	LSTM (Testing)	GRU (Training)	GRU (Testing)
RMSE	24,01	27,54	19,65	26,41
MAE	17,05	21,27	14,09	21,69
MAPE	1,03%	1,09%	0,84%	1,12%
R^2	0,9906	0,7825	0,9937	0,9004
Accuracy	98,97%	98,91%	99,16%	98,88%

3.2. Performance Analysis

This study conducted a comprehensive analysis of the performance of the GRU and LSTM models in predicting the gold exchange rate using various evaluation metrics. In the RMSE comparison, the GRU model showed better performance both on training data (19.65 versus 24.01) and on testing data (26.41 versus 27.54). The more significant difference in training data indicates GRU's superior ability to learn data patterns, while the consistent superiority in testing data indicates better generalization ability for out-of-sample predictions.

In terms of MAE, the GRU model excelled on the training data with a value of 14.09 compared to 17.05 for the LSTM. However, there is an interesting phenomenon in the testing data where the LSTM shows a slightly lower MAE (21.27 versus 21.69). This phenomenon is likely due to the special characteristics of the test data that are better suited to the LSTM architecture. A similar pattern is seen in the MAPE metric, where

GRU performs better on training data (0.84% versus 1.03%), but LSTM slightly outperforms on testing data (1.09% versus 1.12%), albeit by a very small margin. The most significant difference is seen in the R^2 metric, where GRU consistently outperforms LSTM both on training data (0.9937 versus 0.9906) and on testing data (0.9004 versus 0.7825). The large gap in the testing data indicates GRU's much better ability to capture and generalize complex patterns in the gold exchange rate time series. In terms of accuracy (100% - MAPE), GRU performed better on the training data (99.16% versus 98.97%), while LSTM slightly outperformed on the testing data (98.91% versus 98.88%), albeit with a minimal difference of 0.03%.

Overall, both models demonstrated very high accuracy (above 98%) in predicting the gold exchange rate, but the GRU model showed superiority in most evaluation metrics, especially in the R^2 of the test data which indicates better generalization ability. These results suggest that the simpler GRU architecture may be more effective for predictive modeling of gold exchange rates compared to the more complex LSTM architecture. The superior performance of the hybrid LSTM model in this study can be explained by several interrelated factors.

3.3. Discussio

Based on the evaluation results, the GRU model demonstrated superior performance on the training data across all evaluation metrics used. This phenomenon indicates that the simpler GRU architecture is more effective in identifying and learning patterns in the historical gold exchange rate data. Interestingly, on the test data, there was a variation in performance where the GRU model exhibited lower RMSE and significantly higher R^2, while the LSTM model displayed slightly lower MAE and MAPE. This difference underscores the unique characteristics of each model in predicting new, never-before-seen values. The notable difference in the R^2 values in the testing data (0.9004 for GRU versus 0.7825 for LSTM) provides substantive evidence that the GRU model has better generalizability to predict the gold exchange rate beyond the training data. In addition, although both models experienced an increase in RMSE and MAE values when applied from the training data to the testing data, the increase for the GRU model was relatively smaller for RMSE, indicating greater predictive stability in the context of new data.

The aspect of computational efficiency, although not explicitly measured in this study, also needs to be considered. GRU models are generally known to have shorter training times and lower memory requirements thanks to their simpler architecture compared to LSTMs. These findings are consistent with the results of previous studies by Chung et al. (2014) and Jozefowicz et al. (2015) which show that GRU models, despite having a simpler structure, are able to provide equivalent or even superior performance compared to LSTM models in certain contexts. In the case of gold exchange rate prediction, it seems that the simpler structure of GRU is more adaptive to capture complex patterns in financial time series data that are often influenced by various factors and can fluctuate rapidly. It is worth noting that the difference in performance between the two models in some metrics, particularly MAPE and accuracy on test data, is minimal. This indicates that both models are actually very effective in predicting the gold exchange rate, with an unusually high accuracy rate exceeding 98%, which is an excellent achievement in the context of predicting highly volatile financial data.

4. CONCLUSION

This study has compared the performance of LSTM and GRU models in predicting the gold to USD exchange rate using Yahoo Finance data for the period 2017-2023. The evaluation results show that the GRU model displays superior performance in RMSE and R^2 metrics on the test data, with significantly higher R^2 values (0.9004 compared to 0.7825), indicating better generalization ability. Meanwhile, the LSTM model showed slightly lower MAE and MAPE on the test data, albeit with minimal differences. Both models achieved very high accuracy levels above 98% in predicting the gold exchange rate, but the GRU model emerged as the more optimal choice based on a combination of performance on training data and generalizability on testing data. Nevertheless, this study has several limitations, including the use of daily data without considering external factors such as macroeconomic and geopolitical indicators, limited data period, and the possibility of non-optimal architecture configuration. For future research, it is recommended to integrate external factors as additional features, test model performance on various prediction timeframes, explore hybrid architectures that combine the strengths of LSTM and GRU, and develop and evaluate trading strategies based on prediction models in simulations or backtests. The findings have important implications for computational finance practitioners and investors who utilize gold exchange rate predictions in investment decision-making.

REFERENCES

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

[2] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine

translation," *arXiv Prepr. arXiv1406.1078*, 2014.

[3] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv Prepr. arXiv1412.3555*, 2014.

[4] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *International conference on machine learning*, 2015, pp. 2342-2350.

[5] Z. Halim, S. M. Shuhidan, and Z. M. Sanusi, "Corporation Financial Distress Prediction With Deep Learning: Analysis of Public Listed Companies in Malaysia," *Bus. Process Manag. J.*, vol. 27, no. 4, pp. 1163-1178, 2021, doi: 10.1108/bpmj-06-2020-0273.

[6] A. Flores, H. Tito, and D. Centty, "Recurrent Neural Networks for Meteorological Time Series Imputation," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 3, 2020, doi: 10.14569/ijacsa.2020.0110360.

[7] F. Wang, D. Zhang, G. Min, and J. Li, "Reservoir Production Prediction Based on Variational Mode Decomposition and Gated Recurrent Unit Networks," *Ieee Access*, vol. 9, pp. 53317-53325, 2021, doi: 10.1109/access.2021.3070343.

[8] J. Yuan and Y. Tian, "An Intelligent Fault Diagnosis Method Using GRU Neural Network Towards Sequential Data in Dynamic Processes," *Processes*, vol. 7, no. 3, p. 152, 2019, doi: 10.3390/pr7030152.

[9] Z. Yu, Y. Sun, J. Zhang, Y.-A. Zhang, and Z. Liu, "Gated Recurrent Unit Neural Network (GRU) Based on Quantile Regression (QR) Predicts Reservoir Parameters Through Well Logging Data," *Front. Earth Sci.*, vol. 11, 2023, doi: 10.3389/feart.2023.1087385.

[10] A. Inteha, N. Masood, F. Hussain, and I. A. Khan, "A Data Driven Approach for Day Ahead Short Term Load Forecasting," *Ieee Access*, vol. 10, pp. 84227-84243, 2022, doi: 10.1109/access.2022.3197609.

[11] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Short-Term Load Forecasting With Multi-Source Data Using Gated Recurrent Unit Neural Networks," *Energies*, vol. 11, no. 5, p. 1138, 2018, doi: 10.3390/en11051138.

[12] J. Song, G. Xue, Y. Ma, H. Li, Y. Pan, and Z. Hao, "An Indoor Temperature Prediction Framework Based on Hierarchical Attention Gated Recurrent Unit Model for Energy Efficient Buildings," *Ieee Access*, vol. 7, pp. 157268-157283, 2019, doi: 10.1109/access.2019.2950341.

[13] Y. Gao, R. Wang, and E. Zhou, "Stock Prediction Based on Optimized LSTM and GRU Models," *Sci. Program.*, vol. 2021, pp. 1-8, 2021, doi: 10.1155/2021/4055281.

[14] K. He and J. Qian, "Research on Stock Prediction Algorithm Based on CNN and LSTM," *Acad. J. Comput. Inf. Sci*., vol. 5, no. 12, 2022, doi: 10.25236/ajcis.2022.051215.

[15] S. Zhang and F. Wen, "Multifractal Behaviors of Stock Indices and Their Ability to Improve Forecasting in a Volatility Clustering Period," *Entropy*, vol. 23, no. 8, p. 1018, 2021, doi: 10.3390/e23081018.

[16] Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, "Recurrent Neural Networks for Multivariate Time Series With Missing Values," *Sci. Rep.*, vol. 8, no. 1, 2018, doi: 10.1038/s41598-018-24271-9.

[17] S. Park, Y. Seonwoo, J. Kim, J. Kim, and A. Oh, "Denoising Recurrent Neural Networks for Classifying Crash-Related Events," *Ieee Trans. Intell. Transp. Syst*., vol. 21, no. 7, pp. 2906-2917, 2020, doi: 10.1109/tits.2019.2921722.

[18] G. Weiss, Y. Goldberg, and E. Yahav, "On the practical computational power of finite precision RNNs for language recognition," *arXiv Prepr. arXiv1805.04908*, 2018.

[19] D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, 2015, pp. 1422-1432.

[20] M. A. Zaytar and C. El Amrani, "Sequence to sequence weather forecasting with long short-term memory recurrent neural networks," *Int. J. Comput. Appl.*, vol. 143, no. 11, pp. 7-11, 2016.