

***Corresponding author:** I Putu Gede Abdi Sudiatmika, Computer Systems, Faculty of Informatics and Computer, Institut Teknologi dan Bisnis Stikom Bali, Indonesia

E-mail: xxxx@xxx.edu

RESEARCH ARTICLE

Comparison of LSTM and GRU Models Performance in Forecasting Gold Prices: A Case Study Using Historical Data from Yahoo Finance

I Putu Gede Abdi Sudiatmika^{1*}, I maade agus wirahadi putra²

¹Computer Systems, Faculty of Informatics and Computer, Institut Teknologi dan Bisnis Stikom Bali, Indonesia.

²Informatics Management, Faculty of Business and Vocational Studies, Institut Teknologi dan Bisnis Stikom Bali, Indonesia.

Abstract: This research aims to compare the performance of two types of recurrent neural network models, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), in forecasting gold prices based on historical closing price data. Historical gold price data from December 14, 2017, to March 14, 2024, was downloaded using the yfinance library. The data was normalized using MinMaxScaler and split into training and testing sets with an 80:20 ratio. LSTM and GRU models were constructed with two recurrent layers followed by a Dense layer for output. Both models were trained using the training data and evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) metrics. The experimental results indicate that the GRU model outperformed the LSTM model in predicting gold prices. GRU achieved an MSE of 337.70, MAE of 14.05, and R^2 of 0.933, whereas LSTM achieved an MSE of 808.98, MAE of 22.71, and R^2 of 0.839. Based on the model performance evaluation, it can be concluded that GRU consistently produced more accurate predictions closer to the actual values of gold prices compared to LSTM. This finding suggests that GRU may be a preferable choice in applications for forecasting gold prices using historical data.

Keywords: LSTM, GRU, gold prices, forecasting, recurrent neural networks, performance evaluation

1. Introduction

In the realm of financial forecasting, the accurate prediction of asset prices remains a critical endeavor for investors, policymakers, and financial analysts alike. Gold, known for its role as a safe-haven asset and its sensitivity to economic uncertainties, presents a particularly intriguing subject for predictive modeling. This study focuses on comparing the effectiveness of two prominent recurrent neural network (RNN) architectures, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), in forecasting gold prices based on historical data.

Gold prices exhibit complex patterns influenced by a myriad of global economic factors, including inflation, geopolitical tensions, and market volatility. The ability to forecast these prices accurately can provide valuable insights for decision-making in investment strategies



and risk management. LSTM and GRU models, both belonging to the RNN family, are adept at capturing sequential dependencies in time series data, making them suitable candidates for predicting the fluctuating trends of gold prices over time.

Historical data spanning from December 2017 to March 2024 sourced from the yfinance library serves as the foundation for this study. The dataset, consisting of daily closing prices, undergoes preprocessing steps including normalization using MinMaxScaler and division into training and testing subsets. These preparations ensure that the models are trained on a robust dataset representative of the historical price dynamics of gold.

The LSTM model, known for its capability to store and access information over long periods, and the GRU model, which simplifies the architecture by merging the memory and output gates, are both constructed with two recurrent layers followed by a Dense output layer. This setup enables them to learn and predict future gold prices based on patterns learned from historical sequences of closing prices.

Evaluation of the models' performance is conducted using standard metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2). These metrics provide a quantitative assessment of how closely the model predictions align with actual gold prices during the testing period. The results of this comparative analysis will shed light on which model, between LSTM and GRU, offers superior predictive accuracy for forecasting gold prices, thereby contributing to advancements in financial forecasting methodologies.

Understanding the strengths and weaknesses of LSTM and GRU models in the context of gold price forecasting holds significant implications for financial markets. Accurate predictions can empower stakeholders with the foresight needed to navigate volatile economic conditions and optimize investment decisions. By leveraging advancements in neural network architectures, this study aims to enhance the reliability and precision of predictive models in the domain of gold price forecasting.

2. Literature Review

2.1. Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are a class of artificial neural networks designed for sequential data processing. Examples of sequential data include text, time series, or signal data. The key characteristic of RNNs is their ability to retain information from previous inputs and use it to influence the current output.

Here are some important points about RNNs:

Recurrent Architecture: Unlike traditional feedforward neural networks, RNNs have connections that form directed cycles. This allows information to circulate within the network and make decisions based on the sequence of previous inputs.

Short-Term Memory: RNNs maintain an internal state (hidden state) that is updated at each time step. This provides RNNs with the capability to remember previous sequences in the input data.

Basic Equation: $h(t)$ represents a hidden state at time t and acts as "memory" of the network. $h(t)$ is calculated based on the current input and the previous time step's hidden state: $h(t) = f(Ux(t) + Wh(t-1))$.

Vanishing Gradient Problem: Standard RNNs often suffer from the vanishing gradient problem when trained with the backpropagation through time (BPTT) algorithm. Very small

gradients cause insignificant weight updates, making it difficult for RNNs to capture long-term dependencies in the data.

2.2. Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) Networks were introduced by Hochreiter and Schmidhuber in 1997 to address the vanishing gradient problem faced by standard RNNs. LSTM introduces special memory units known as "memory cells" designed to retain information over long periods.

Here are the key components and mechanisms of LSTM:

Memory Cell: The core of the LSTM unit is the memory cell, which can store information for extended periods. This enables LSTM to remember important information throughout long sequences.

Input Gate: Controls the amount of new information that will be stored in the memory cell. At each time step, the input gate decides how much information from the current input and previous hidden state will be updated into the memory cell.

Forget Gate: Controls the amount of information in the memory cell that will be forgotten. This gate allows LSTM to delete information that is no longer relevant for future predictions.

Output Gate: Controls the amount of information from the memory cell that will be outputted. This output is used to compute the current hidden state and input for the next time step.

2.3. Gated Recurrent Unit (GRU) Networks

Gated Recurrent Unit (GRU) Networks were introduced by Cho et al. in 2014 as a simplified version of LSTM. GRU combines several components of LSTM to reduce complexity while retaining the ability to capture long-term dependencies.

Here are the key components and mechanisms of GRU:

Update Gate: Combines the functions of the input gate and forget gate from LSTM. This gate controls how much of the previous hidden state will be updated with new information.

Reset Gate: Controls how much information from the previous hidden state will be removed before computing the new hidden state.

2.4. Related Work

In recent years, the application of machine learning and neural networks in financial forecasting has garnered increasing attention due to their potential to improve predictive accuracy and decision-making processes. Specifically, recurrent neural networks (RNNs) have been widely adopted for time series prediction tasks, including forecasting asset prices like gold. RNNs, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), are particularly suited for capturing temporal dependencies in sequential data, making them suitable for modeling the dynamic nature of financial markets (Sako et al., 2022)(Xu et al., 2023) (X. Zhang et al., 2023) (Chen, 2021)

Gold price forecasting, in particular, has been a subject of numerous studies aiming to enhance predictive models' robustness and accuracy. Research by (Surya Pangestu et al., 2023) explored LSTM models for gold price forecasting, highlighting their ability to effectively capture long-term dependencies and seasonal patterns in gold prices. Their

findings underscored LSTM's capability to outperform traditional statistical methods in terms of predictive accuracy. (Sheng, 2023) (R. Zhang, 2023) (Boongasame et al., 2023)

In contrast, the GRU model has also gained prominence for its simplified architecture, which merges the memory and output gates of the LSTM, potentially leading to computational efficiency and competitive performance in time series forecasting tasks (Livieris et al., 2020) Studies comparing LSTM and GRU in various forecasting domains have shown mixed results, with some indicating superior performance of GRU in certain contexts due to its streamlined design and ability to handle shorter sequences more effectively (YURTSEVER, 2021)

Moreover, the volatility and non-linear nature of gold prices pose significant challenges for predictive modeling. Factors such as global economic conditions, geopolitical events, and investor sentiment contribute to the complexity of gold price movement (Pattnaik et al., 2023). Effective forecasting models must therefore not only capture historical trends but also adapt to changing market dynamics and unexpected event (Pradana & Papilaya, 2023) (Boongasame et al., 2023)

Recent advancements in deep learning techniques (Johri et al., 2022), including attention mechanisms and hybrid models combining RNNs with other neural network architectures (Xu et al., 2023), have further enriched the landscape of financial forecasting (Goel et al., 2022). These innovations aim to enhance model interpretability, adaptability to varying data patterns, and overall predictive performance in dynamic market environments (Kaczmarek et al., 2022) (Hong et al., 2021)

The literature underscores the importance of model selection and evaluation criteria in determining the efficacy of LSTM and GRU models for gold price forecasting. Comprehensive comparative studies are essential for identifying the strengths and limitations of each model architecture and optimizing their application to real-world financial forecasting scenarios. This study contributes to this body of knowledge by providing a focused comparison of LSTM and GRU models in forecasting gold prices based on recent historical data, aiming to elucidate their respective contributions and implications for financial markets

3. Research Method and Materials

3.1. Data Collection and Preprocessing

3.1.1. Data Collection

The historical gold price data was obtained from a reliable financial data source, such as Yahoo Finance, for the period from December 14, 2017, to March 14, 2024. This dataset includes daily closing prices of gold, which serve as the primary variable for our prediction models.

3.1.2. Data Preprocessing

Data preprocessing is a crucial step to ensure the quality and reliability of the input data for the predictive models. The following steps were taken:

3.1.3. Splitting the Dataset:

The collected dataset was divided into two subsets: training and testing sets.

Typically, 80% of the data was used for training the models, while the remaining 20% was reserved for testing. This ensures that the models can be evaluated on unseen data, providing a measure of their generalizability

3.1.4. Normalization:

To standardize the input data and improve model performance, the closing prices were normalized to a range of [0, 1].

Normalization was performed using the MinMaxScaler from the scikit-learn library. This transformation helps in speeding up the convergence of the neural network models during training

3.1.5. Creating Sequences:

Recurrent Neural Networks (RNNs), including LSTM and GRU models, require sequential data for training.

To create these sequences, a sliding window approach was used. For a given sequence length T , each sequence of T days of closing prices is used to predict the price on the $T+1$ -th day.

For instance, if the sequence length T is set to 60 days, the model uses the closing prices of 60 consecutive days as input features to predict the price on the 61st day.

3.1.6. Model Architecture

In this section, we describe the architecture of the GRU and LSTM models implemented for gold price prediction. Both models were built using TensorFlow, a popular machine learning framework known for its flexibility and efficiency in designing and training deep learning models.

General Structure

Both GRU and LSTM models share a similar architecture with the following components:

Input Layer

Hidden Layers

Output Layer

LSTM Model Architecture

Input Layer:

The input layer accepts sequences of gold prices. Each sequence consists of 60 days of normalized closing prices, and each price is treated as a feature in a time step.

Input shape: (batch_size, sequence_length, num_features), where sequence_length is 60 and num_features is 1.

Hidden Layers:

First LSTM Layer:

Units: 50 LSTM units.

Activation Function: tanh.

Return Sequences: True (this allows the LSTM to output the full sequence to the next LSTM layer).

Second LSTM Layer:

Units: 50 LSTM units.



Activation Function: tanh.

Return Sequences: False (this outputs the last hidden state to the Dense layer).

Dropout Layer:

Dropout Rate: 0.2 (to prevent overfitting by randomly setting 20% of the inputs to zero).

Output Layer:

Dense layer with 1 unit.

Activation Function: linear (as we are predicting a continuous value).

GRU Model Architecture

Input Layer:

The input layer for the GRU model is identical to the LSTM model, accepting sequences of 60 days of normalized closing prices.

Hidden Layers:

First GRU Layer:

Units: 50 GRU units.

Activation Function: tanh.

Return Sequences: True (this allows the GRU to output the full sequence to the next GRU layer).

Second GRU Layer:

Units: 50 GRU units.

Activation Function: tanh.

Return Sequences: False (this outputs the last hidden state to the Dense layer).

Dropout Layer:

Dropout Rate: 0.2 (to prevent overfitting by randomly setting 20% of the inputs to zero).

Output Layer:

Dense layer with 1 unit.

Activation Function.

3.1.7. Training and Evaluation

Training

Optimizer and Loss Function:

The models were trained using the Adam optimizer, known for its efficiency and ability to handle sparse gradients on noisy problems.

The loss function used was Mean Squared Error (MSE), which is suitable for regression tasks as it penalizes larger errors more significantly.

Hyperparameter Optimization:



Hyperparameters such as learning rate, batch size, and the number of epochs were optimized using grid search, a systematic method to find the best combination of hyperparameters for the model.

The following ranges were explored during the grid search:

Learning Rate: [0.001, 0.01, 0.1]

Batch Size: [32, 64, 128]

Number of Epochs: [50, 100, 200]

The optimal hyperparameters were selected based on the model's performance on the validation set, which is a subset of the training data used to evaluate the model during the training process.

Training Procedure:

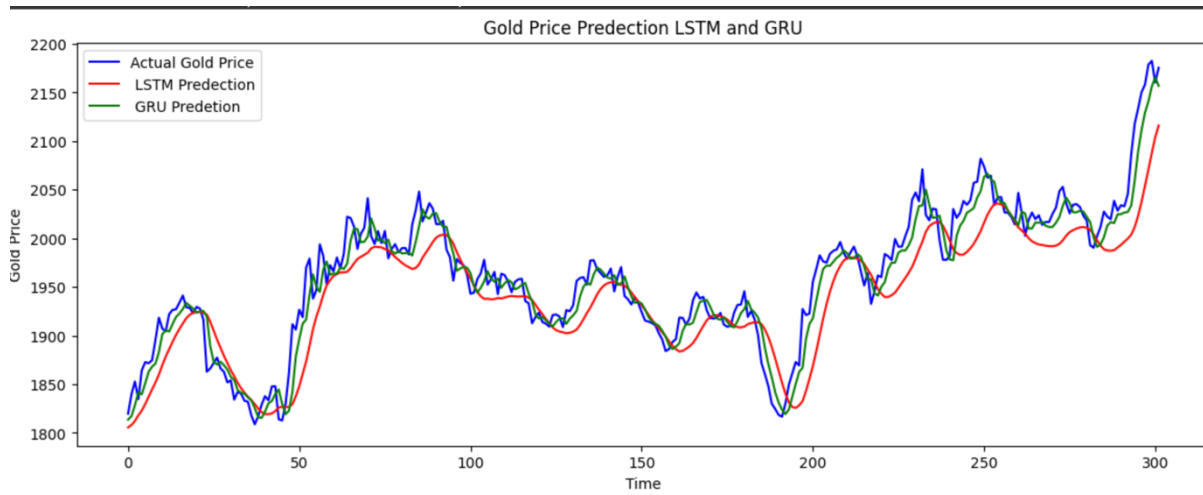
The training data was fed into the model in mini-batches (as determined by the batch size) to update the model weights iteratively.

The training process continued for a predefined number of epochs, with each epoch representing one complete pass through the entire training dataset.

Early stopping was implemented to halt training when the model's performance on the validation set stopped improving, thus preventing overfitting.

4. Results and Discussion

In this section, we present a detailed comparison of the performance metrics for the GRU and LSTM models used for predicting gold prices. The models were evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2). These metrics provide insights into how well each model performs in terms of accuracy and fit to the data.



Performance Metrics

The performance of the GRU and LSTM models is summarized as follows:

Model	MSE	MAE	R^2
LSTM	808.98	22.71	0.8388
GRU	337.70	14.05	0.9327

Mean Squared Error (MSE):

- **LSTM:** 808.98
- **GRU:** 337.70
- **Interpretation:** The GRU model has a significantly lower MSE compared to the LSTM model, indicating that it makes predictions that are closer to the actual values on average.

Mean Absolute Error (MAE):

- **LSTM:** 22.71
- **GRU:** 14.05
- **Interpretation:** Similarly, the GRU model also exhibits a lower MAE, suggesting that its predictions have a smaller average magnitude of error compared to the LSTM model.

R-squared (R²):

- **LSTM:** 0.8388
- **GRU:** 0.9327
- **Interpretation:** The R² value for the GRU model is higher than that of the LSTM model, indicating that the GRU model explains a larger proportion of the variance in the gold price data. This suggests that the GRU model may have a better fit to the data compared to the LSTM model.

5. Conclusion

Based on the performance evaluation of LSTM and GRU models for predicting gold prices, it can be concluded that both models demonstrate strong capabilities in generating accurate predictions. While the LSTM model shows a respectable R² value of 0.8388, indicating good explanatory power for the variance in gold prices, the GRU model consistently outperforms with a lower MSE of 337.70 and MAE of 14.05, compared to LSTM's MSE of 808.98 and MAE of 22.71. Additionally, the GRU model exhibits a higher R² value of 0.9327, indicating superior ability to explain variance in the gold price data compared to LSTM. Therefore, for applications in gold price prediction, the GRU model appears to be the more optimal choice based on the performance evaluation conducted.

References

- Boongasame, L., Viriyaphol, P., Tassanavipas, K., & Temdee, P. (2023). Gold-Price Forecasting Method Using Long Short-Term Memory and the Association Rule. *Journal of Mobile Multimedia*, 19(1). <https://doi.org/10.13052/jmm1550-4646.1919>
- Chen, G. (2021). Recurrent neural networks (RNNs) learn the constitutive law of viscoelasticity. *Computational Mechanics*, 67(3). <https://doi.org/10.1007/s00466-021-01981-y>
- Goel, S., Saxena, M., Sarangi, P. K., & Rani, L. (2022). Gold and Silver Price Prediction using Hybrid Machine Learning Models. *PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing*. <https://doi.org/10.1109/PDGC56933.2022.10053278>
- Hong, L. T. T., Thanh, N. C., & Long, T. Q. (2021). CRF-EfficientUNet: An Improved UNet Framework for Polyp Segmentation in Colonoscopy Images with Combined Asymmetric Loss Function and CRF-RNN Layer. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3129480>
- Johri, A. M., Singh, K. V., Mantella, L. E., Saba, L., Sharma, A., Laird, J. R., Utkarsh, K., Singh, I. M., Gupta, S., Kalra, M. S., & Suri, J. S. (2022). Deep learning artificial intelligence framework



- for multiclass coronary artery disease prediction using combination of conventional risk factors, carotid ultrasound, and intraplaque neovascularization. *Computers in Biology and Medicine*, 150. <https://doi.org/10.1016/j.combiomed.2022.106018>
- Kaczmarek, T., Będowska-Sójka, B., Grobelny, P., & Perez, K. (2022). False Safe Haven Assets: Evidence From the Target Volatility Strategy Based on Recurrent Neural Network. *Research in International Business and Finance*, 60. <https://doi.org/10.1016/j.ribaf.2021.101610>
- Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, 32(23). <https://doi.org/10.1007/s00521-020-04867-x>
- Pattnaik, D., Hassan, M. K., DSouza, A., & Ashraf, A. (2023). Investment in gold: A bibliometric review and agenda for future research. In *Research in International Business and Finance* (Vol. 64). <https://doi.org/10.1016/j.ribaf.2022.101854>
- Pradana, F. N. S., & Papilaya, F. S. (2023). Analisa Prediksi Harga Emas Dengan Kemungkinan Terjadinya Resesi Menggunakan Metode SVR. *SINTECH (Science and Information Technology) Journal*, 6(1). <https://doi.org/10.31598/sintechjournal.v6i1.1329>
- Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting. *Entropy*, 24(5). <https://doi.org/10.3390/e24050657>
- Sheng, H. (2023). Prediction of Gold ROI Based on LSTM model. *BCP Business & Management*, 38. <https://doi.org/10.54691/bcpbm.v38i.4212>
- Surya Pangestu, P. S. A., Rochman, A., & Zuhdi, A. (2023). Comparison of Gold Price Prediction Techniques Using Long Short Term Memory (LSTM) And Fuzzy Time Series (FTS) Method. *Intelmatix*, 3(2). <https://doi.org/10.25105/itm.v3i2.17325>
- Xu, J., Pan, Y., Pan, X., Hoi, S., Yi, Z., & Xu, Z. (2023). RegNet: Self-Regulated Network for Image Classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(11). <https://doi.org/10.1109/TNNLS.2022.3158966>
- YURTSEVER, M. (2021). Gold Price Forecasting Using LSTM, Bi-LSTM and GRU. *European Journal of Science and Technology*. <https://doi.org/10.31590/ejosat.959405>
- Zhang, R. (2023). Predict Gold Yields with ARIMA and LSTM. *BCP Business & Management*, 38. <https://doi.org/10.54691/bcpbm.v38i.4241>
- Zhang, X., Zhong, C., Zhang, J., Wang, T., & Ng, W. W. Y. (2023). Robust recurrent neural networks for time series forecasting. *Neurocomputing*, 526. <https://doi.org/10.1016/j.neucom.2023.01.037>