

Comparison of LSTM and GRU Models for Forex Prediction

Mohammad Rezza Pahlevi^{1)*}, Kusrini²⁾, Tonny Hidayat³⁾

^{1,2,3)} Universitas Amikom Yogyakarta, Indonesia

¹⁾ pahlevirezza48@students.amikom.ac.id, ²⁾ kusrini@amikom.ac.id, ³⁾ tonny@amikom.ac.id

Submitted : Jul 17, 2023 | **Accepted** : Jul 20, 2023 | **Published** : Oct 1, 2023

Abstract: Trading foreign currencies worth trillions of dollars takes place daily in the forex market, characterized by highly volatile movements. The forex market operates on bid and ask prices, with exchange rates determined by the principles of supply and demand. Trading involves currency pairs like EUR/USD, where the value of the Euro is compared to the US Dollar, serving as a basis for analyzing price fluctuations. Due to the volatile nature of forex, market participants must make informed decisions when buying and selling, as improper choices can result in financial losses. One approach to mitigating risk in forex trading decisions is through the use of forecasting techniques. This research study employs LSTM and GRU methods to predict forex trends, which are evaluated using various dataset divisions. The most accurate results are obtained using a dataset of 4979, split into three equal parts: 80% for training, 10% for validation, and 10% for testing. This approach yields an RMSE value of 0.054, MAPE of 0.037, and R-square of 97%.

Keywords: Forex, Prediction, LSTM, GRU, RMSE, MAPE

INTRODUCTION

The global foreign exchange market, known as forex, holds the title for being the largest currency exchange market worldwide. It facilitates the trading of foreign currencies, with trillions of dollars being exchanged daily. Due to the nature of foreign currency trading, the forex market experiences significant fluctuations in value. Unlike traditional stock markets, the forex market operates around the clock, allowing trading at any time. However, it adheres to the Australian, Asian, Euro, and North American time zones as its principal time zones, each with its specific opening and closing hours. Forex trading revolves around bid and ask prices, determining the buying and selling rates. Foreign exchange rates are determined by the market based on the principles of supply and demand with its specific opening and closing hours. Forex trading revolves around bid and ask prices, determining the buying and selling rates. Foreign exchange rates are determined by the market based on the principles of supply and demand. Although forex shares similarities with stock trading, it possesses certain distinctions (Hu et al., 2018).

Given the market's high volatility, non-linearity, and chaos, the currency market is among the most challenging. The currency market is particularly volatile since it is not governed by a single institution or organization (Henriquez & Kristjanpoller, 2019). For the past several decades, experts have been primarily interested in forecasting the currency market. Leverage is one of the most significant market-related tools. The market for foreign exchange does not require enormous sums of money, unlike typical markets like the stock market. Leverage, in its simplest form, is the capacity to open trades on any currency pair with just rudimentary capital protection (Islam & Hossain, 2021).

Trading currency pairs such as EUR / USD is a comparison of the value of the Euro currency against the Dollar as a basis for researching, rising and falling currency prices in forex move volatily, so a market participant must be able to decide positions in buying and selling. Because improper decisions can lead to losses. One way to reduce risk in making decisions in buying and selling in forex can use forecasting

*Mohammad Rezza Pahlevi



The LSTM model is the greatest option for forex forecasting since it includes state cells, which allow it to store data for a long time. Like most repeating neural networks, the drawback of LSTM is its lengthy training period and complicated parameter combination (Muflikhah & Cholissodin, 2022). To improve the accuracy of the GRU forecast process for forex, the autocorrelation function and partial autocorrelation function will be used to identify significant delays. (Mahan Zaky, 2022).

According to the preceding definition, the goal of this study is to assess how well algorithms work in predicting forex prices by focusing on a number of supporting factors or other factors connected to how they affect the value of forex prices themselves. Processing time series data typically involves the application of the LSTM and GRU models. The use of LSTM and GRU to forecasting EUR/USD prices and evaluating the outcomes of such forecasting will be the main emphasis of this study. The purpose of this study is to shed light on the application of LSTM and GRU in forecasting forex trading.

LITERATURE REVIEW

Research on deep learning's application to predicting. This study's literature review examined studies that employed the LSTM and GRU models. The study's result, weakness, and comparison are laid up in Table 1.

Tabel 1 Literature Review Matrix Using LSTM and GRU

Author	Result	Weakness	Comparison
(Qi et al., 2020)	Create a prediction system that enables precise trading tactics with less risk. The most optimal model for EUR/GBP data at 15-minute intervals reached RMSE 1.5×10^3 with MAPE 0.12%.	The study only used experimental data with 15-minute intervals without comparing them to longer time intervals.	This study will use predictions using longer time intervals.
(Sarangi et al., 2020)	The model used to make predictions uses 2 models, namely ANN and ANN-GA which produce RMSE 0.39 and 0.018930.	Researchers used one experiment in conducting evaluations, namely with RMSE.	This study will add more currencies to test the methods used such as MAPE and R square
(Ulina et al., 2020)	Model used LSTM and CEEMDAN-LSTM using hidden layer 32 and batch size with a dropout value of 0.04. Yields RMSE values of 0.009546 and MAPE 0.612705.	The study conducted an experiment using dropout 0.04	The study conducted an experiment using dropout 0.05 to produce better scores.
(Cabrera, 2019)	The best results in 3 models are ARIMA with MSE results of 0.00024910 with USDCNY currency.	The study only used one scenario testing and training data, namely 70% and 30%	In this study will conduct several scenarios in testing data testing and training,
(Abedin et al., 2021)	The results of the test with 80:20 data sharing the best results were found in SGD / USD with an RMSE value of 0.0025.	Researchers only use RMSEs in evaluating the models used	This research will add model evaluation metrics such as MAPE and R square.

*Mohammad Rezza Pahlevi

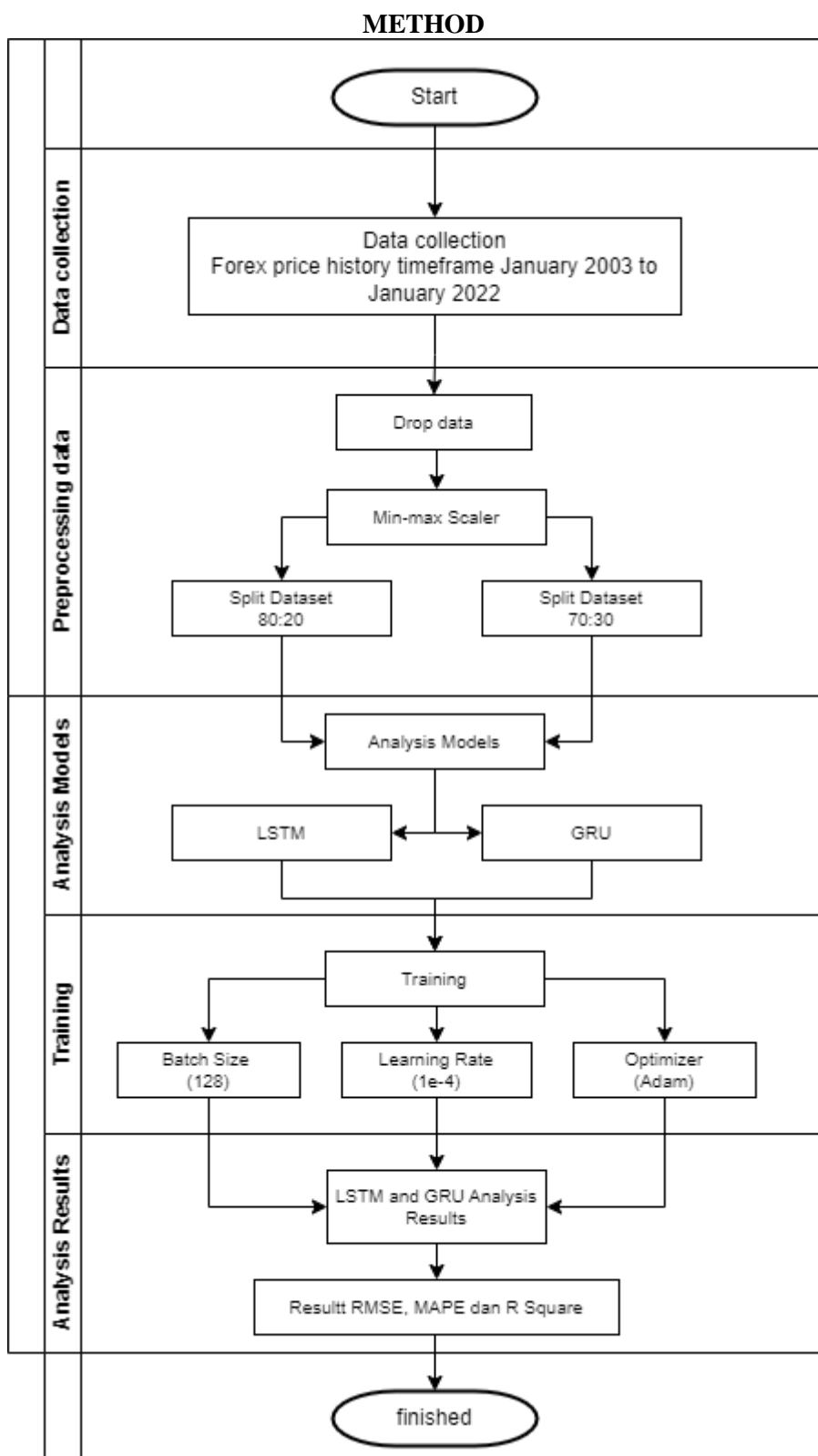


Figure 1. Research Flow

Figure 1 depicts the direction of the study that will be done. This research is divided into five stages. The following is a description of the steps of research mentioned above in order:

*Mohammad Rezza Pahlevi



Dataset Collection

In this study, we use a primary data obtained from the global financial market news website, Investing.com. Dataset with a time span of January 2003 to January 2022 with a total of 4979 data used with a daily data timeline. The data used in making predictions is EURUSD using the average price of the price on that day. The data used in this study is backward because in this study it will produce a distance between the data used and the prediction results. The collected data will be processed into several timelines that will be used in testing and divided into several data timelines that will produce maximum prediction value.

Preprocessing

Researchers do preprocessing with *Google Collabs* by checking data, calculating dataset data. The application of hardware libraries will be applied to this model because hard modularity and flexibility make it easy to build complex LSTM and GRU architectures by combining different layers and activity functions. Features include hardware libraries such as, regularization, dropouts, initializers, optimizers, and more. This allows researchers to experiment with various configurations and techniques to improve the performance of LSTM models.

Dataset Allocation

In this study, two test scenarios are run by partitioning the dataset into three parts and applying various percentages to each component. Table 2 lists the scenario test case.

Tabel 2. Dataset Allocation

No	Training	Testing	Validation
1	70%	15%	15%
2	80%	10%	10%

The EURUSD dataset will be used for the data training procedure at this time. Two of the models utilized in this work, the LSTM and GRU models, underwent training. The prediction process then uses the output of each model by loading the model file. Predictive values will be derived from each model's output using the data that was designated for testing and validation. In order to determine the projected and evaluation of the model performance outcomes, a denormalization technique was also used to the test data.

Model Selection

This study makes currency forecasts using LSTM and GRU models.

Long Short-Term Memory

Application in the field of deep learning with LSTM is one of the methods that is often used in prediction. LSTM is a popular model with the power of handling unknown size gaps between signal in data noise. Created the LSTM. LSTM is universal so that when enough network units are available, any data is calculated by the computer assuming when it has a properly calibrated weight matrix (Puspita, 2022). A memory-cell-containing RNN version is the LSTM. The four additional sub-neurons that make up each LSTM neuron serve as memory cells. These sub-neurons' weights aid the LSTM in remembering lengthy sequences. Figure 2 depicts the first LSTM cell, that has an output gate and an input gate.

*Mohammad Rezza Pahlevi



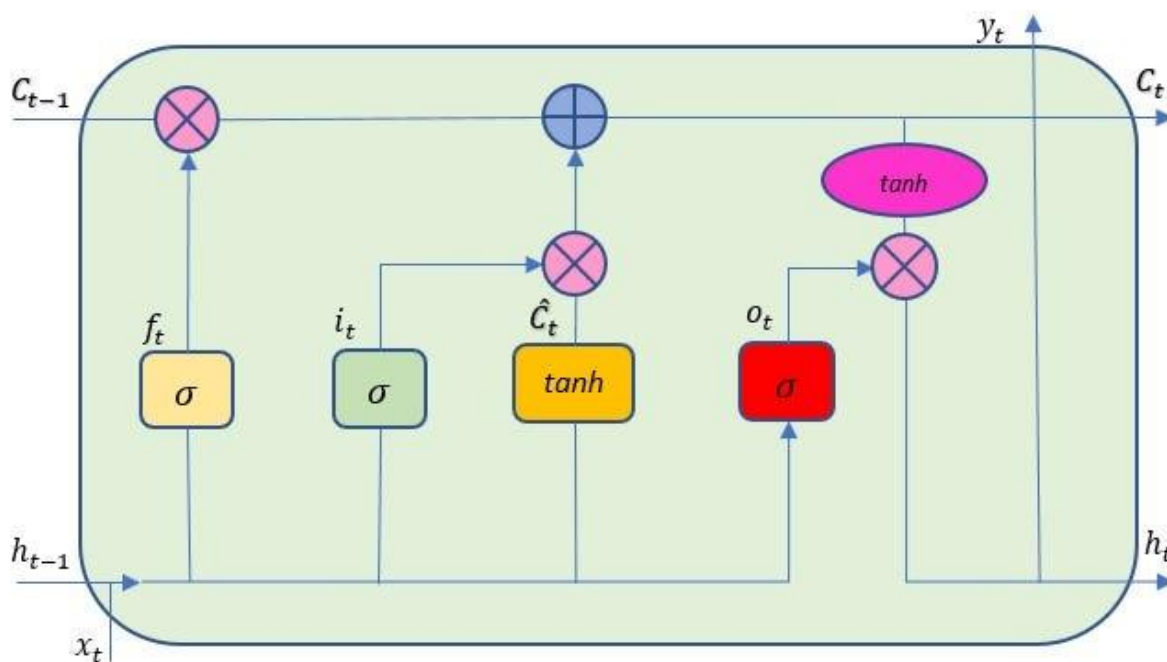


Figure 2 LSTM Architecture

The following formula is used to execute computations in the LSTM computing process:

$$f_t = \sigma(Wf * [ht-1, Xt] + bf) \quad (1)$$

$$i_t = \sigma(Wi * [ht-1, Xt] + bi) \quad (2)$$

$$\hat{C}_t = \tanh(Wc * [ht-1, Xt] + bc) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (4)$$

$$o_t = \sigma(Wo * [ht-1, Xt] + bo) \quad (5)$$

The memory cell state value (t) in LSTM, which regulates each neuron's memory, as well as the input gate (it), output gate (ot), and forget gate (ft) in each LSTM neuron. It is a filtered version of the current cell state (C_t), the forget gate determines what information to discard from the previous cell state (C_{t-1}), takes the previous cell state and the current input (X_t) allow the LSTM to learn which data to keep and which data to discard.(Primananda & Isa, 2021).

Gated Recurrent Unit

The Gated Recurrent Unit (GRU) is a branch of the RNN. GRU is the application of RNN for forecasting cases. The data output is an integer or discrete number, while for regression cases the data output is a real or continuous number. Using the concept of insensitive loss function, introduced by Vapnik, GRU can be generalized to perform function approximation or regression approaches.

Suppose we have a training data set, (x_i, y_i) , $i = 1, \dots$, with input data $x = (x_1, x_2, \dots, x_n)$ and corresponding output $y = \{y_1, \dots, y_n\} \subseteq \mathbb{R}$. With GRU, we want to find a function $f(x)$ that has the greatest divergence over all training data from the real target y_i . We get a perfect regression when the value is equal to 0. Suppose we have a function as a regression line in the equation (Mahan Zaky, 2022)

Training Model

The training model uses 500 epochs, a learning rate of 0.0001, Adam as the optimizer, and a batch size of 128 in each test run on each pre-trained model.

Analysis Result

At analysis result, an evaluation of the utilized model is conducted to identify the errors generated during the prediction process. The measurement of model error from predictions made using quantitative data is called RMSE. RMSE is used to determine the size of the distribution of data point deviations from linear regression lines or to determine the concentration of data around linear regression

*Mohammad Rezza Pahlevi

lines (Satyo Bayangkari Karno et al., 2020). This method is called RMSE where the smaller the RMSE value (the closer to 0), the more accurate the measurement results will be (Dhanardono Bhima & Atmodjo Danang, 2022).

$$RMSE = \sqrt{(\sum(y_i - \hat{y}_i)^2 / n)} \quad (6)$$

y_i indicates the target variable's real (true) value for the i -th data point, \hat{y}_i reflects the target variable's anticipated value for the i -th data point, \sum denotes the summation, i.e., summing over all data points, n the amount of data points in the dataset as a whole.

Mean Absolute Percentage Error (MAPE) is the average absolute percentage error, what is meant by mean absolute percentage error is a statistical measure of the accuracy of forecasting in the forecasting process. MAPE measurement can be used by the general public because MAPE is easy to understand and use to predict forecast accuracy. The Mean Absolute Percentage Error method provides information on how well the forecast error compares to the actual value of the circuit (Khrisna Wardhani Anindya & Israwan Fajar, 2022). To calculate the overall total MAPE is made by first subtracting the actual data value from the forecast data, then dividing by the actual data and multiplying by 100, then dividing by the amount of data present.

$$MAPE = (\sum | (y_i - \hat{y}_i) / y_i |) * (100 / n) \quad (7)$$

y_i represents the true (actual) value of the target variable for the i -th data point, \hat{y}_i represents the predicted value of the target variable for the i -th data point, $| |$ denotes the absolute value, ensuring the differences are always positive, \sum denotes the summation, i.e., summing over all data points, and is the total number of data points in the dataset.

R-Squared (or coefficient of determination) is a measure of statistical model evaluation that assesses the goodness of a regression model. It helps data analysts to explain performance models compared to basic models. Its value lies between 0 and 1. A value close to 0 represents a bad model while a value near 1 represents a perfect match (Fandango et al., 2021). Sometimes, R-squared produces negative values. This means your model is worse than the average base model. We can describe R-squared using the following formula:

$$R^2 = 1 - (SSR / SST) \quad (8)$$

R^2 is the coefficient of determination, which shows the percentage of the dependent variable's variation that can be predicted from the independent variables using the model. SSR (Sum of Squared Residuals) is the sum of the squared differences between the predicted values (\hat{y}_i) and the actual values (y_i). SST (Total Sum of Squares) is the sum of the squared differences between the actual values (y_i) and the mean of the dependent variable (\bar{y}).

RESULT

The Global Financial Market News website provided primary data for the data collection process employed in this study Investing.com. The dataset consists of 2 types of datasets, namely datasets with a period of January 2003 to January 2022 with a total of 4979 data used with a daily data timeline. The data used in making predictions is EURUSD using the average price of the price on that day. The data used in this study is backward because in this study it will produce a distance between the data used and the prediction results. The collected data will be processed into several timelines that will be used in testing and divided into several data timelines that will produce maximum prediction value. Based on the amount of data taken is shown in Tabel 3.

*Mohammad Rezza Pahlevi



Date	Price	Open	High	Low	Change
2003-01-01	1.0492	1.0494	1.0502	1.0492	-0.09%
2003-01-02	1.0362	1.0492	1.0498	1.0334	-1.24%
2003-01-03	1.0423	1.0359	1.0438	1.0436	-0.59%
2003-01-06	1.0469	1.0417	1.0497	1.0414	-0.44%
2003-01-07	1.0416	1.0463	1.0468	1.0394	-0.51%
2022-01-25	1.1299	1.1324	1.1330	1.1263	-0.21%
2022-01-26	1.1237	1.1301	1.1312	1.1235	-0.55%
2022-01-27	1.1143	1.1240	1.1244	1.1132	-0.84%
2022-01-28	1.1143	1.1146	1.1174	1.1121	-0.00%
2022-01-31	1.1233	1.1147	1.1249	1.1249	-0.81%

Tabel 3 Datasets from 2003 to 2022

The distribution of training and testing data had a direct impact on prediction accuracy in a number of tests. A data allotment of 70/15/15 and 80/10/10 is employed. Table 3 displays the models that were tried during modeling for this investigation. Comparing the estimation results with the actual values yields a performance rating for each of the models in Table 4 that are provided. The GRU model comes in second, with LSTM producing the best RMSE, MAPE, and R square values.

Tabel 4. Analysis Results

Model	Data Allocation	RMSE	MAPE	R square
LSTM	70/15/15	0.058	0.040	97%
	80/10/10	0.055	0.037	96%
GRU	70/15/15	0.058	0.041	96%
	80/10/10	0.054	0.037	97%

Python programming language and Google collaborative tools are used to create the LSTM model. The deep learning library Keras is used to create LSTM models. Table 5 lists the hyperparameters that were specified during the validation procedure.

Table 5. Hyperparameters Values

No	Hyperparameters	Information
1	Layers	2
2	Optimizer	Adam
3	Loss	MSE
4	learning_rate	0.0001
5	Epoch	500

*Mohammad Rezza Pahlevi



Two layers make up the LSTM model used for this research: an input layer with one layer with 50 neurons per layer, and an output layer made up of two solid layers with 25 neurons each. The batch size and epoch values for the used LSTM model are 128 and 500, respectively. The MSE value was used as the loss value in the study, and the Adam Optimizer was used to run this model. When trained in the LSTM layer, solid layers are utilized to record changes or patterns of movement in specific gold values. Results of an LSTM 80 percent of the dataset is used for training, 10 percent is used for testing, and 10 percent is used for validation.

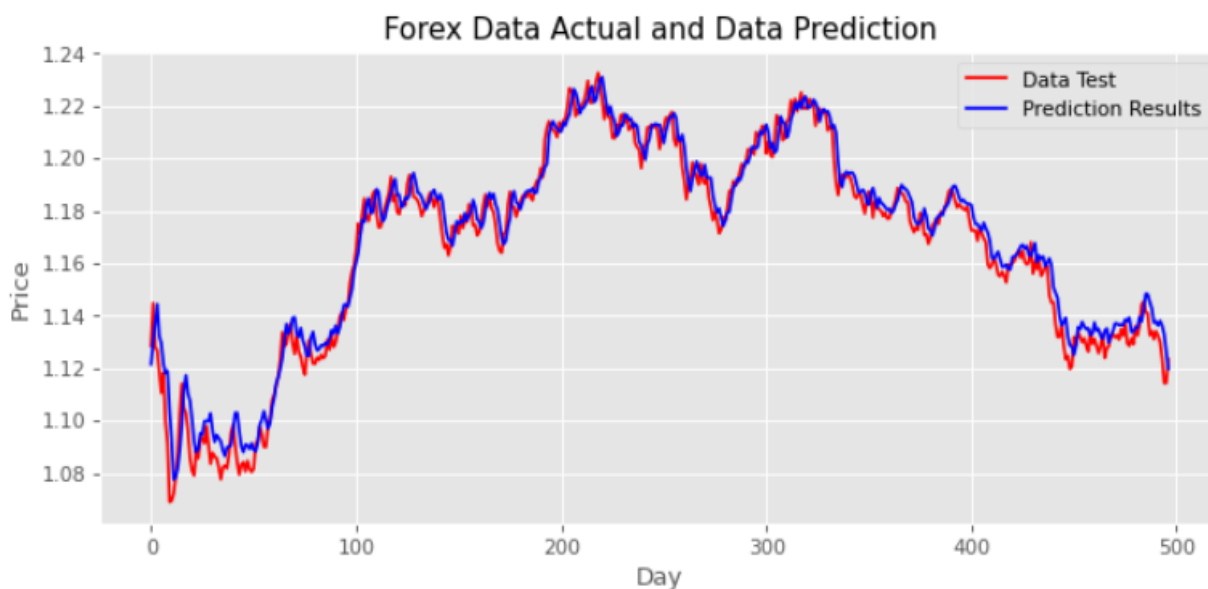


Figure 4. Prediction with LSTM Models

Additionally, Python programming language and Google collaborative tools are used for GRU model training. Using the Keras deep learning package, GRU models are created. Table 6 lists the hyperparameters that were specified during the validation procedure.

Table 6. Hyperparameters Values

No	Jenis	Informasi
1	Layers	2
2	Optimizer	Adam
3	Loss	MSE
4	learning_rate	0.0001
5	Epoch	500

The GRU model created for this study contains two layers: an input layer made up of one LSTM layer with 50 neurons per layer, and an output layer made up of two solid layers with 25 neurons each. The LSTM model utilized has an epoch value of 500 and a batch size value of 128. The study employed the Adam Optimizer to execute this model, using MSE value as the loss value. When trained in the LSTM layer, solid layers are utilized to record changes or patterns of movement in specific gold values. The outcomes of an LSTM model with an allocation of 80% of the dataset for training data, 10% for testing data, and 10% validation.

*Mohammad Rezza Pahlevi



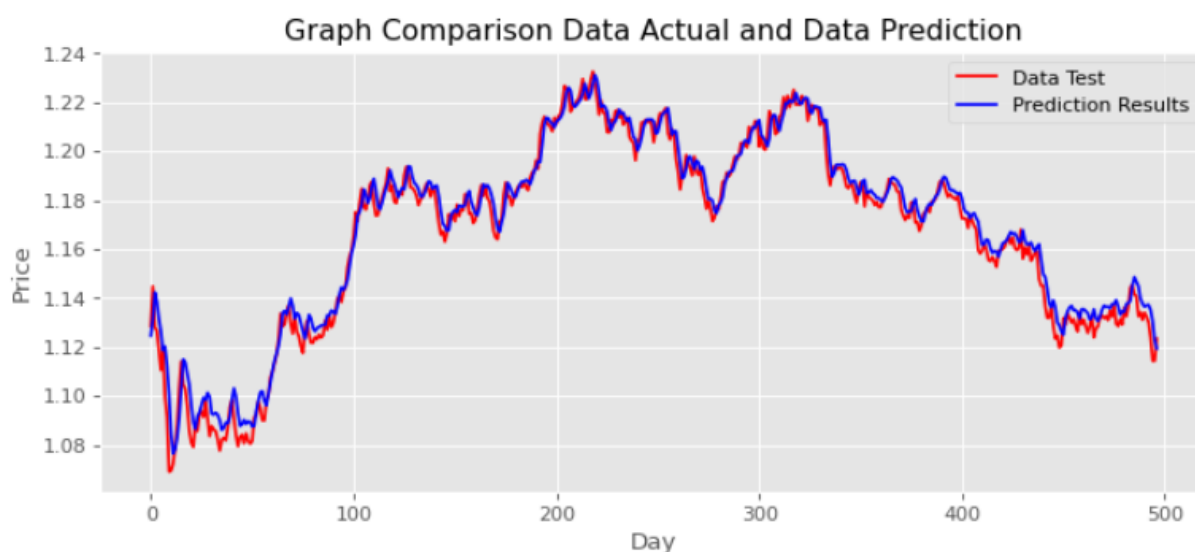


Figure 5. Prediction with GRU Model

DISCUSSIONS

The study's performance metrics, which were used to compare the effectiveness of the LSTM and GRU models, are shown in Table 4. A daily data timeline is used when using data to forecast forex. Predictions are made using EURUSD data and that day's average price. R square evaluates how well the regression model fits the data, and RMSE measures predicting accuracy, where the predicting accuracy increases with decreasing forecasting RMSE.

In both models, an 80% training data, 10% testing data, and 10% validation data split produces the best RMSE and MAPE values. According to the study's findings, the LSTM model with 50 neurons, 128 batches, and 500 epochs has an RMSE of 0.055 and the GRU model with 50 neurons, 128 batches, and 500 epochs has an RMSE of 0.054. Without a doubt, the performance above shows that GRU performs better when applying the RMSE logic. Both the LSTM and GRU models had the same MAPE result of 0.037, with the smaller the value, the better the result, according to MAPE, which estimates the average percentage of error. Our study only used daily data so that it could extend a longer research period and use weekly data samples. This study used a longer time interval than previous studies conducted by (Qi et al., 2020) By applying daily time intervals, by increasing the use of dropouts to produce better values from the research done (Ulina et al., 2020). In this study added the results of analysis in the form of MAPE and Rsquare to compare the anasalis results of the research (Abedin et al., 2021). LSTM and GRU models, Along with several models from previous studies, there are different training and data allocation tests that directly affect prediction outputs. The results demonstrate that the GRU model outperforms the LSTM model, with an RMSE value of 0.054, MAPE of 0.037, R square of 97%, and data allocation of 80/10/10 for the GRU model.

CONCLUSION

In this work, we used daily time series data to assess how well LSTM and GRU models predicted forex. Daily EURUSD data from January 2003 to January 2022 make up the study's dataset. Both models' performance is gauged using RMSE, MAPE, and R square to determine accuracy. When actual forex is compared to the predicted value of each model, it can be seen that GRU performs better than LSTM. The GRU model, which uses an 80/10/10 data allocation with 50 neurons per batch and an epoch of 500, has an RMSE value of 0.054, MAPE of 0.037, and R square of 97%.

*Mohammad Rezza Pahlevi



REFERENCES

- Abedin, M. Z., Moon, M. H., Hassan, M. K., & Hajek, P. (2021). Deep learning-based exchange rate prediction during the COVID-19 pandemic. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-021-04420-6>
- Cabrera, F. (2019). *Application of ARIMA, RNN and LSTM on Foreign Exchange Rate Prediction*.
- Dhanardono Bhima, & Atmodjo Danang. (2022). *Pemodelan Tiga Dimensi Menggunakan Total Station, Terrestrial Laser Scanner Dan Unmanned Aerial Vehicle*. Deepublish,.
- Fandango, Armando., Idris, Ivan., & Navlani, Avinash. (2021). *Python Data Analysis - Third Edition*. Packt Publishing.
- Henríquez, J., & Kristjanpoller, W. (2019). A combined Independent Component Analysis–Neural Network model for forecasting exchange rate variation. *Applied Soft Computing Journal*, 83. <https://doi.org/10.1016/j.asoc.2019.105654>
- Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T. Y. (2018). Listening to chaotic whispers: A deep learning framework for news-oriented Stock trend prediction. *WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining, 2018-February*, 261–269. <https://doi.org/10.1145/3159652.3159690>
- Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. *Soft Computing Letters*, 3, 100009. <https://doi.org/10.1016/j.socl.2020.100009>
- Khrisna Wardhani Anindya, & Israwan Fajar. (2022). *Buku di Google Play Teknik Peramalan Pada Teknologi Informasi*. Get Press.
- Mahan Zaky, A. (2022). Forecasting Jumlah Penggunaan Obat Digestive Enzymes dengan Algoritma GRU. *Journal of Informatics and Vocational Education*, 5(2), 48–55. <https://jurnal.uns.ac.id/joive/index>
- Muflikhah, L., & Cholissodin, I. (2022). *Peramalan Kasus Positif COVID-19 di Jawa Timur menggunakan Metode Hybrid ARIMA-LSTM* (Vol. 6, Issue 9). <http://j-ptiik.ub.ac.id>
- Primananda, S. B., & Isa, S. M. (2021). Forecasting Gold Price in Rupiah using Multivariate Analysis with LSTM and GRU Neural Networks. *Advances in Science, Technology and Engineering Systems Journal*, 6(2), 245–253. <https://doi.org/10.25046/aj060227>
- Puspita, H. (2022). *Pengantar Teknologi Informasi*.
- Qi, L., Khushi, M., & Poon, J. (2020, December 16). Event-Driven LSTM for Forex Price Prediction. *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020*. <https://doi.org/10.1109/CSDE50874.2020.9411540>
- Sarangi, P. K., Chawla, M., Ghosh, P., Singh, S., & Singh, P. K. (2020). FOREX trend analysis using machine learning techniques: INR vs USD currency exchange rate using ANN-GA hybrid approach. *Materials Today: Proceedings*, 49, 3170–3176. <https://doi.org/10.1016/j.matpr.2020.10.960>
- Satyo Bayangkari Karno, A., Noer Ali, J. K., & Bekasi, K. (2020). Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long ShortTerm Memory). *Journal of Information and Information Security (JIFORTY)*, 1(1), 1–8. <http://ejurnal.ubharajaya.ac.id/index.php/jiforty>
- Ulina, M., Purba, R., & Halim, A. (2020, November 3). Foreign Exchange Prediction using CEEMDAN and Improved FA-LSTM. *2020 5th International Conference on Informatics and Computing, ICIC 2020*. <https://doi.org/10.1109/ICIC50835.2020.9288615>