Foreign Exchange Rates Prediction Based on Comparative Models

Yuyao Wang

Department of AIAC, Xi'an Jiaotong-liverpool University, Suzhou, China

Yuyao.Wang22@student.xjtlu.edu.cn

Abstract. With the development of globalization, fluctuations in the foreign exchange market have the huge potential to influence global economy. Machine Learning technology shows great potential in the field of financial forecasting than traditional statistical methods in recent years. This research selected statistical model Autoregressive Integrated Moving Average (ARIMA) and machine learning models Linear Regression and Long- and Short-Term Memory (LSTM) to make training and testing of historical data on the forex market. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-Squared were chosen as evaluation indexes to evaluate the performance of each model and compare their performance in forex exchange rate forecasts. The results show that ARIMA and LSTM could make better predictions than Linear Regression, between which LSTM has higher degree of fitting data. This research provides a detailed analysis comparing the performance of different models in forex exchange rate forecasting whose results are beneficial to guide forex market forecast in practical application and provide reference for future research.

Keywords: Statistics, Machine Learning, ARIMA, Linear Regression, LSTM.

1. Introduction

The foreign exchange (forex) market stands as one of the most dynamic and crucial components of the global financial system, whose many complicated factors have great impacts on foreign exchange rates [1]. Foreign exchange rates play a crucial role in international trade, investment and economic globalization. The fluctuations of foreign exchange rates could bring profound influence to the economic system and market participants. Moreover, foreign exchange rate fluctuations may affect the investment decisions of multinational corporations and domestic inflation or deflation. Especially, sharp fluctuations in foreign exchange rates have great impacts on the stability of financial markets, increasing investors' worries of market risk [2]. Therefore, researchers have gradually begun to predict foreign exchange rates by using statistical models and machine learning models in order to help economic decision-making and policies formulation with the development of innovative technologies, controlling market risks [1].

Statistical methods and machine learning models have been widely applied in healthcare, education and finance to solve different problems related to data analysis [3-5], among which predictions of foreign exchange rates has always been a popular topic. Colombo & Pelagatti, 2020 have made predictions on foreign exchange rates through statistical learning models, which performed well in short term forecasting. What's more, they ranked the importance of different variables and made analysis of the correlation between the different variables and results in order to explain how statistical models operated [6]. Apart from this, Dautel et al., 2020 used different models of machine learning to forecast foreign exchange rates, who stated that deep learning has had huge impacts in finance field. Therefore, four models --- long short-term memory networks, gated recurrent units, traditional recurrent networks and feedforward networks were used to carry on the research. They concluded that a simple neural network may have better performance than a complex deep neural network in some respects [7]. Besides this, there is one issue that whether statistical models or deep learning models perform better. Through the research conducted by Aggarwal & Sahani, 2020 and Bangyuan, n.d., it is obvious that deep learning models perform better on both long and short term predictions on foreign exchange rates than statistical learning models, making great contribution to financial researches [1, 8].

This paper aims to predict foreign exchange rates. More specifically, two foreign exchange rates were chosen: exchanges rates of JPY to USD and exchange rates of GBP to USD in recent 5 years as data set. Considering the data is time series data, this paper will use 3 models: Autoregressive Integrated Moving Average (ARIMA), Linear Regression, and Long Short Term Memory (LSTM) to predict foreign exchange rates.

2. Methodology

2.1. Data Description

This research selected two foreign exchange rates: exchange rate of JPY to USD and exchange rate of GBP to USD in recent 5 years from February 22, 2019 to February 22, 2024 from the website: https://cn.investing.com/currencies [9] as data set. It is obvious that these data are time series data. The predictive objective is the change of foreign exchange rate data. The basic information of the data set is shown in Table1, which has 1502 rows and 2 columns. According to the data set, the trend of original data is shown in Fig. 1 and Fig. 2.

Table 1. Data Information

Column	Non-Null Count	Dtype
Exchange Rate of JPY to USD	1305 non-null	Float64
Exchange Rate of GBP to USD	1305 non-null	Float64

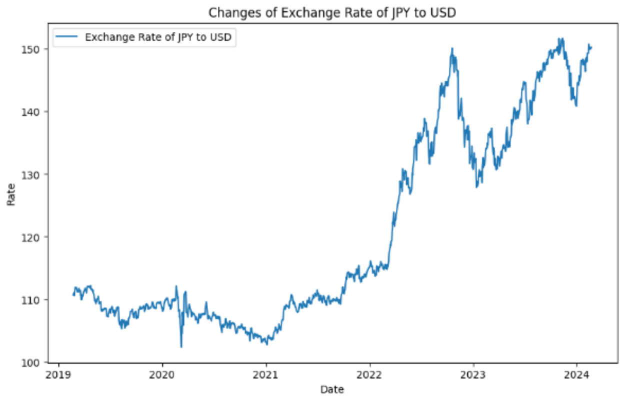


Figure 1. Trend of JPY to USD (Photo/Picture credit: Original)

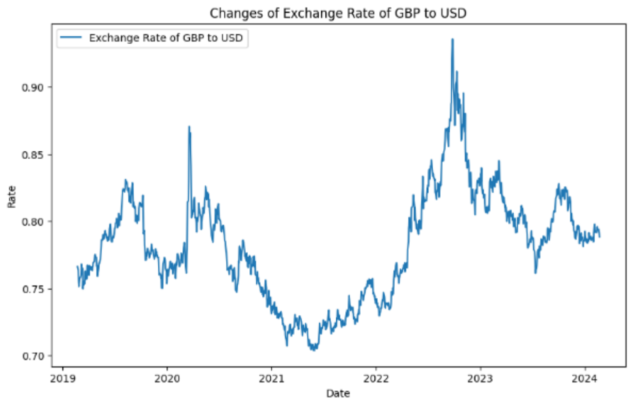


Figure 2. Trend of GBP to USD (Picture credit: Original)

2.2. Data Pre-Processing

Before making the fitting and predicting of the model, the research preprocessed the time series data from the data set. Firstly, the 'Date' column is set as the index column. Secondly, this research drawn the box plot of the two rates respectively (shown in Fig. 3), from which the outliers in the column 'Exchange Rate of GBP to USD' are found. In order to deal with the outliers, the research calculated the value of quartile: Q1 (25th percentile) = 0.7518, Q3 (75th percentile) = 0.8094, IQR (Interquartile Range) = 0.0576, the lower bound is 0.6654, and the upper bound is 0.8958. Then this research replaced the outliers with the median of column ‘Exchange Rate of GBP to USD’ and got the processed data (shown in Fig. 4).

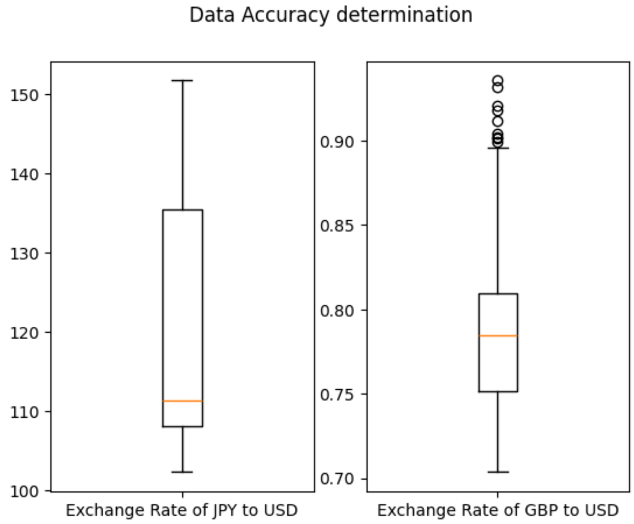


Figure 3. Box Plot (Picture credit: Original)

```
DataFrame with outliers replaced by median:
            Exchange Rate of JPY to USD  Exchange Rate of GBP to USD  Outlier
Date
2019-02-22                       110.69                       0.7661    False
2019-02-25                       111.06                       0.7635    False
2019-02-26                       110.58                       0.7547    False
2019-02-27                       111.00                       0.7513    False
2019-02-28                       111.39                       0.7540    False
...                                 ...                          ...      ...
2024-02-16                       150.21                       0.7935    False
2024-02-19                       150.11                       0.7939    False
2024-02-20                       149.99                       0.7920    False
2024-02-21                       150.28                       0.7910    False
2024-02-22                       150.27                       0.7883    False

[1305 rows x 3 columns]
```

Figure 4. Processed Data (Picture credit: Original)

Furthermore, this research Use 'MinMaxScaler' from the ‘sklearn' to standardize the time series data and scale the data into a specific range [-1, 1]. Next, this research set the time window size ‘lookback’ to 100 and split the data into training and test sets. By iterating the time series data, the previous' lookback 'time step is taken as the input feature, and the value of the next time step is taken as the target value. After the above processing, the data is transformed into a format suitable for LSTM model training, and important time information is retained, which helps to improve the model's interpretability and generalization ability.

2.3. ARIMA

ARIMA represents ‘Autoregressive Integrated Moving Average’ which is made up of autoregression (AR), difference (I), and moving average (MA). This model could capture both linear and nonlinear relationships of data, which is a common method to analyze time series data, which is a common statistical model used to predict changes in time series. The workflow of ARIMA is shown in Fig. 5 [10].

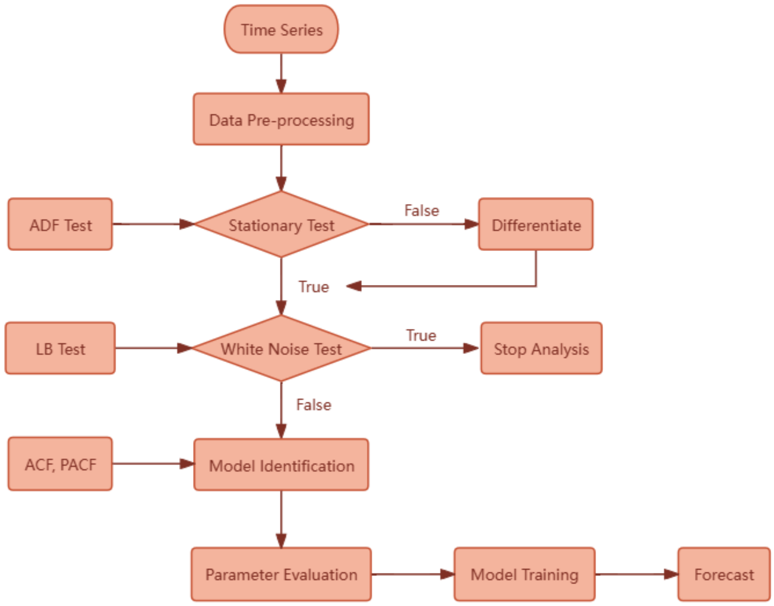


Figure 5. ARIMA Workflow (Picture credit: Original)

Taking the 'Exchange Rate of JPY to USD' as example, the research first used ADF test to test the stationarity of time series data and found that the p-value was greater than the significance level (usually 0.05), the null hypothesis could not be rejected, that is, the data was not stationary. Therefore, the research made difference on the data, and ADF test is conducted on the column data after two differences, and the p-value is less than 0.05, so the data passes the stationarity test. Then, LB test was used to conduct white noise test on the data, and it was judged that it was not a white noise sequence, so it could continue to be analyzed. ACF and PACF charts (shown in Fig. 6 and Fig. 7) were drawn to estimate the values of p and q. The last 100 data were selected from the whole column of data, and 90% were divided into training sets and 10% into test sets. Furthermore, the research tried to find the optimal ARIMA model parameters through iteration, which were finally determined as p=10, d=1, q=3. The parameters of 'Exchange Rate of GBP to USD' were determined by the same operation as p=12, d=3, q=2.

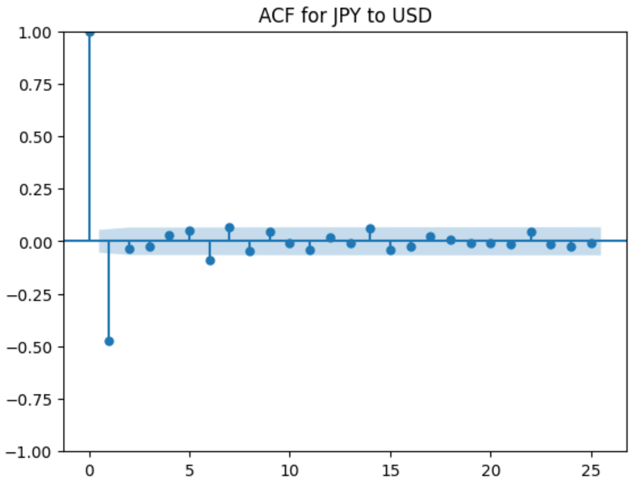


Figure 6. ACF for JPY to USD (Picture credit: Original)

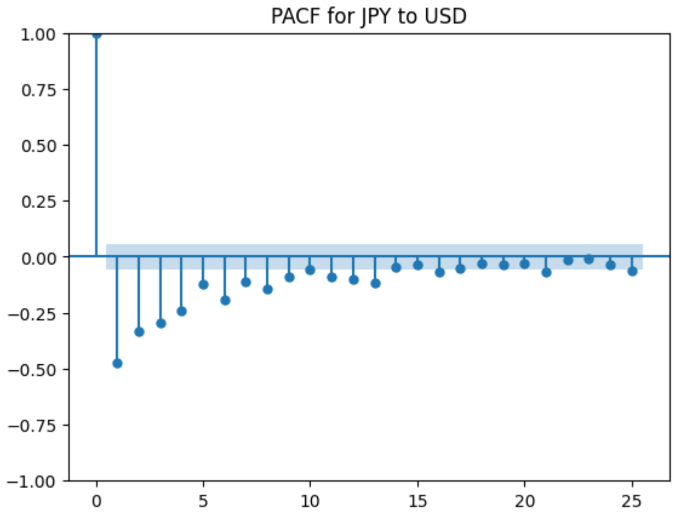


Figure 7. PACF for GBP to USD (Picture credit: Original)

2.4. Linear Regression

Linear Regression is a simple statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed again. The workflow of linear regression is shown in Fig. 8. This research put ‘Exchange Rate of JPY to USD’ and ‘Exchange Rate of GBP to USD’ respectively into the linear regression model for fitting training [11].

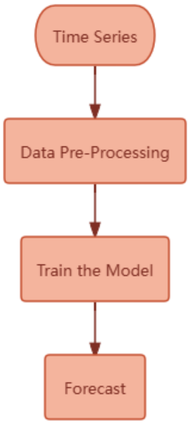


Figure 8. Linear Regression Workflow (Picture credit: Original)

2.5. LSTM

LSTM, also called ‘Long-short Term Memory’, is a type of Recurrent Neural Network (RNN) architecture, specifically designed to address vanishing gradient problem in traditional RNNs, and to capture long-term dependencies in sequential data. The workflow of LSTM is shown in Fig. 9 [12].

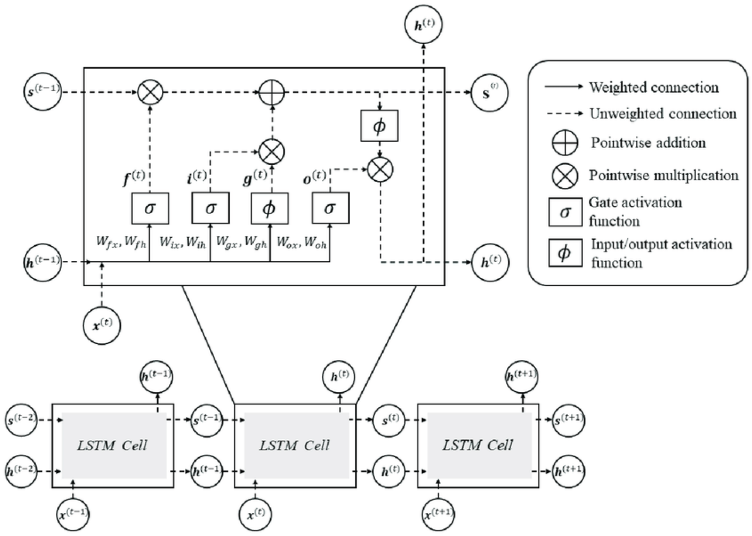


Figure 9. LSTM Workflow [13]

The neural network model used in this research adopts the structure of 2-layer LSTM, whose input dimension is set to 1 which means the number of features per time step is 1, the hidden dimension is set to 32, and the output dimension is 1. During the training process, the model was trained for 200 rounds. This research used the LSTM layer and the fully connected layer, where the LSTM is responsible for processing the long-term dependency of the time series data, and the fully connected layer maps the output of the LSTM layer to the final output dimension. Furthermore, the mean square error Loss function (MSE Loss) was used as the loss function of the model, Adam was used as the optimizer of the model, and the learning rate was set at 0.05. Over 200 training cycles, the research trained the training set repeatedly and recorded the loss value for each training cycle.

3. Results and Discussion

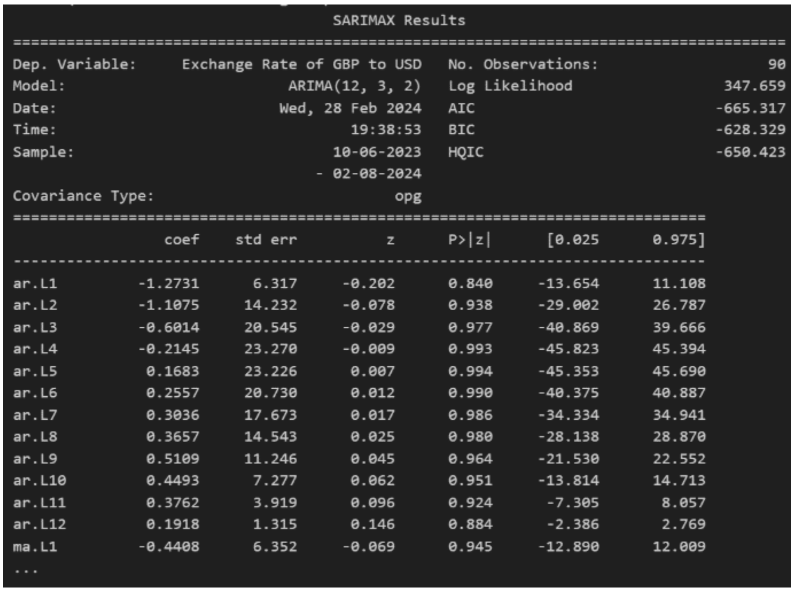
3.1. Results

3.1.1 ARIMA

The results of ARIMA were shown in Fig. 10 to Fig. 13.

```
                                    SARIMAX Results
==========================================================================================
Dep. Variable:     Exchange Rate of JPY to USD   No. Observations:                     90
Model:                          ARIMA(10, 1, 3)   Log Likelihood                  -104.388
Date:                          Wed, 28 Feb 2024   AIC                              236.776
Time:                                  19:24:11   BIC                              271.617
Sample:                              10-06-2023   HQIC                             250.819
                                   - 02-08-2024
Covariance Type:                            opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.4931      0.789      0.625      0.532      -1.054       2.040
ar.L2          0.2945      1.183      0.249      0.803      -2.023       2.612
ar.L3         -0.9410      0.944     -0.997      0.319      -2.791       0.909
ar.L4          0.4211      0.234      1.800      0.072      -0.037       0.880
ar.L5         -0.2061      0.275     -0.749      0.454      -0.746       0.333
ar.L6         -0.1818      0.405     -0.449      0.654      -0.976       0.612
ar.L7          0.2829      0.290      0.974      0.330      -0.286       0.852
ar.L8         -0.1313      0.174     -0.755      0.451      -0.472       0.210
ar.L9          0.0552      0.203      0.272      0.785      -0.342       0.453
ar.L10         0.2240      0.207      1.083      0.279      -0.181       0.629
ma.L1         -0.5298      0.807     -0.657      0.511      -2.111       1.051
ma.L2         -0.4997      1.221     -0.409      0.682      -2.893       1.894
ma.L3          0.9576      0.936      1.023      0.306      -0.877       2.793
...
===================================================================================
```

Figure 10. Results of JPY to USD (Photo/Picture credit: Original)



```
                              SARIMAX Results
==============================================================================
Dep. Variable:     Exchange Rate of GBP to USD   No. Observations:          90
Model:                        ARIMA(12, 3, 2)   Log Likelihood         347.659
Date:                        Wed, 28 Feb 2024   AIC                   -665.317
Time:                                19:38:53   BIC                   -628.329
Sample:                            10-06-2023   HQIC                  -650.423
                                 - 02-08-2024
Covariance Type:                          opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -1.2731      6.317     -0.202      0.840     -13.654      11.108
ar.L2         -1.1075     14.232     -0.078      0.938     -29.002      26.787
ar.L3         -0.6014     20.545     -0.029      0.977     -40.869      39.666
ar.L4         -0.2145     23.270     -0.009      0.993     -45.823      45.394
ar.L5          0.1683     23.226      0.007      0.994     -45.353      45.690
ar.L6          0.2557     20.730      0.012      0.990     -40.375      40.887
ar.L7          0.3036     17.673      0.017      0.986     -34.334      34.941
ar.L8          0.3657     14.543      0.025      0.980     -28.138      28.870
ar.L9          0.5109     11.246      0.045      0.964     -21.530      22.552
ar.L10         0.4493      7.277      0.062      0.951     -13.814      14.713
ar.L11         0.3762      3.919      0.096      0.924      -7.305       8.057
ar.L12         0.1918      1.315      0.146      0.884      -2.386       2.769
ma.L1         -0.4408      6.352     -0.069      0.945     -12.890      12.009
...
```

Figure 11. Results of GBP to USD (Photo/Picture credit: Original)

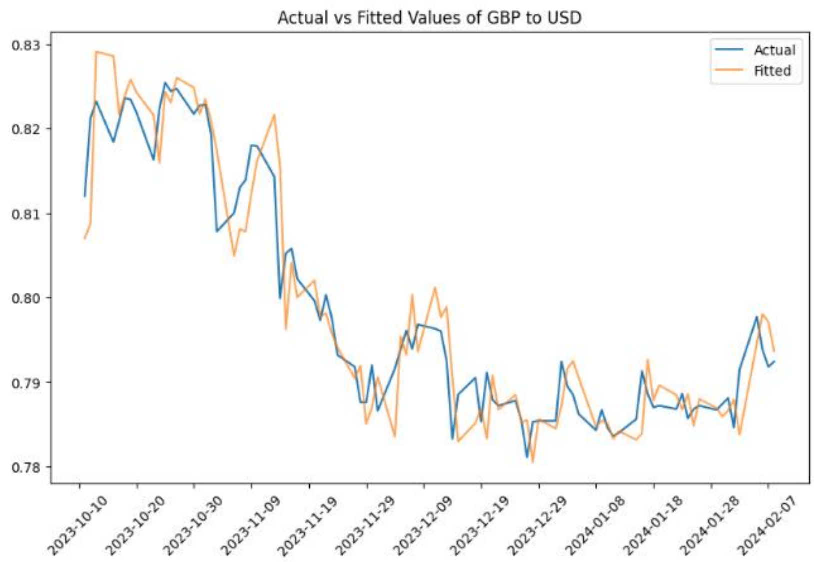


Figure 12. Predictions of train set (GBP to USD) (Photo/Picture credit: Original)

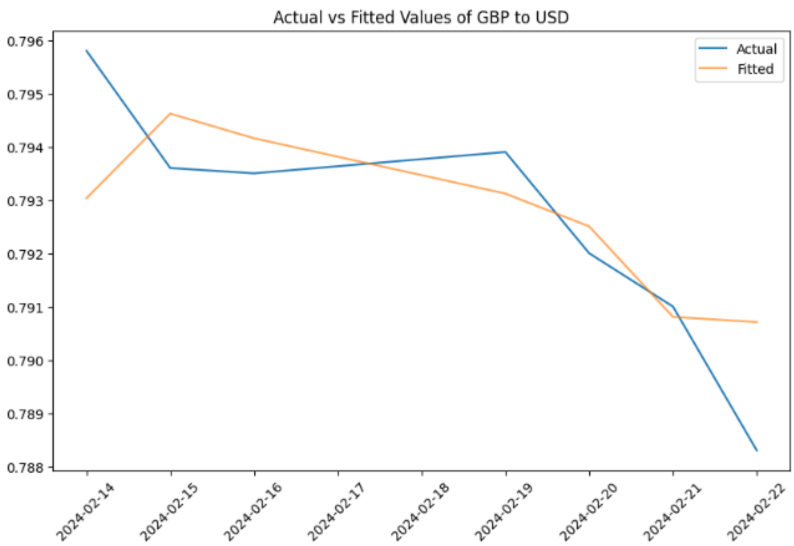


Figure 13. Predictions of test set (GBP to USD) (Photo/Picture credit: Original)

ARIMA model has the advantages of simplicity and interpretability. The architecture of the model is straightforward, making it easy to grasp and execute. However, the processes involved are complex, and it demands strict conditions for time series data. ARIMA might not yield optimal results for datasets characterized by nonlinearity, lack of stationarity, or non-Gaussian distributions.

3.1.2 Linear Regression

The two comparison graphs between the actual value and the predicted value were drawn and shown in Fig. 14 and Fig. 15.

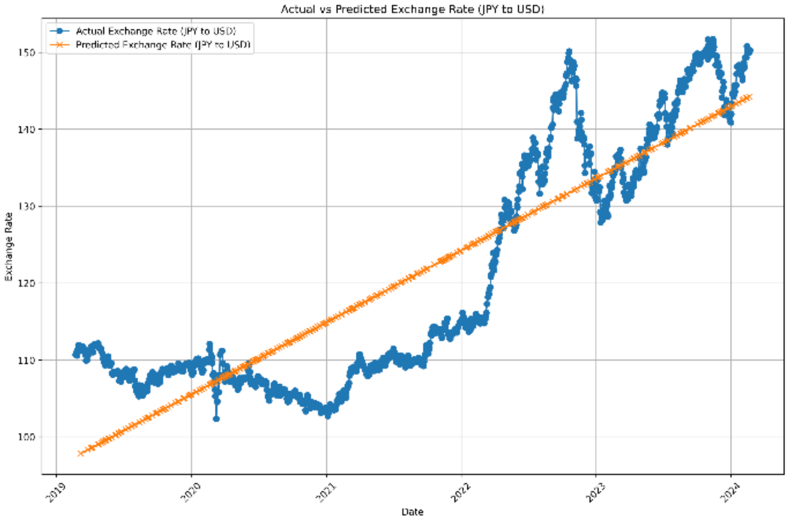


Figure 14. Results of JPY to USD (Photo/Picture credit: Original)

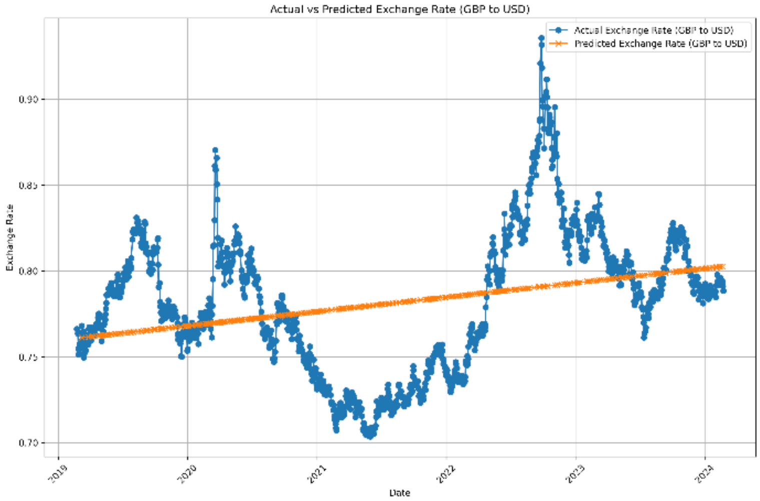


Figure 15. Results of GBP to USD (Photo/Picture credit: Original)

Linear Regression can freely add a variety of independent variables, including time trends, seasonal factors, and other influencing factors, which is easy to understand and implement, so as to better capture the dynamic characteristics of time series data. The parameters of the model are intuitively understandable, yet it falls short in accurately fitting nonlinear data and shows increased sensitivity to outliers.

3.1.3 LSTM

Through the visual analysis of the loss value, the change trend of the loss value in the training process of the model was observed (shown in Fig. 16 and Fig. 17), and the convergence of the model and the training effect were judged.

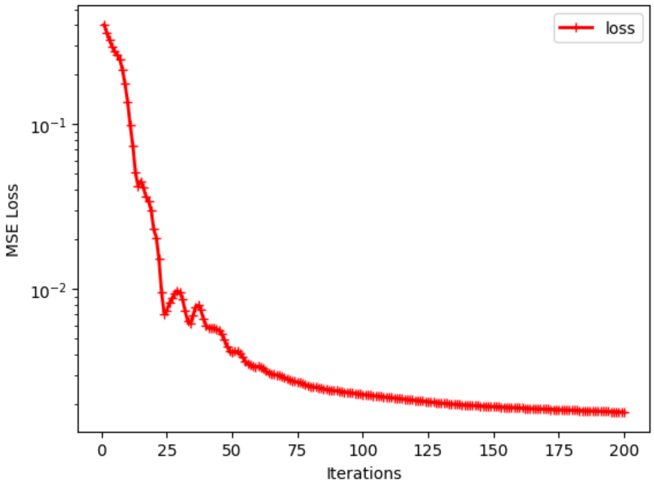


Figure 16. MSE Loss of JPY to USD (Photo/Picture credit: Original)

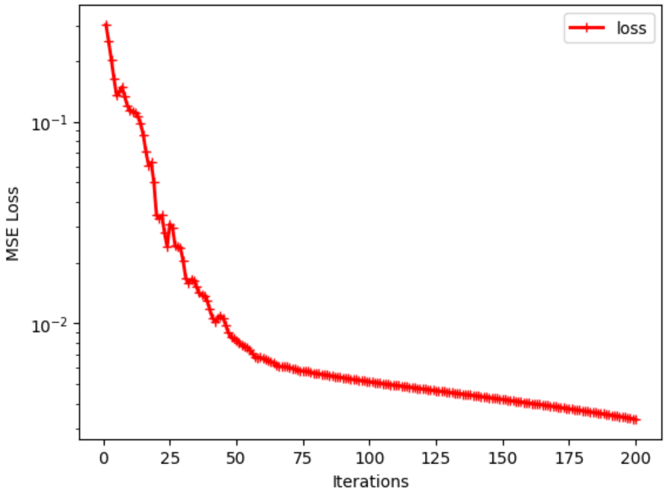


Figure 17. MSE Loss of GBP to USD (Photo/Picture credit: Original)

The comparison result graphs between the actual value and the predicted value by LSTM are shown in Fig. 18 and Fig. 19.

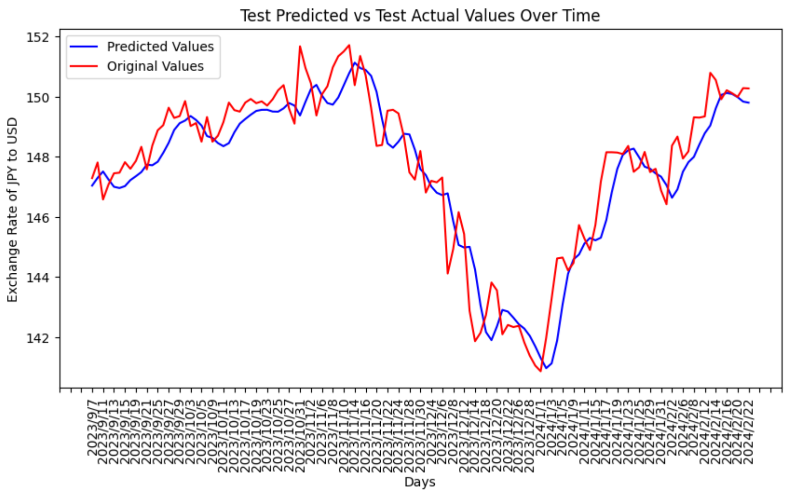


Figure 18. Results of JPY to USD (Photo/Picture credit: Original)

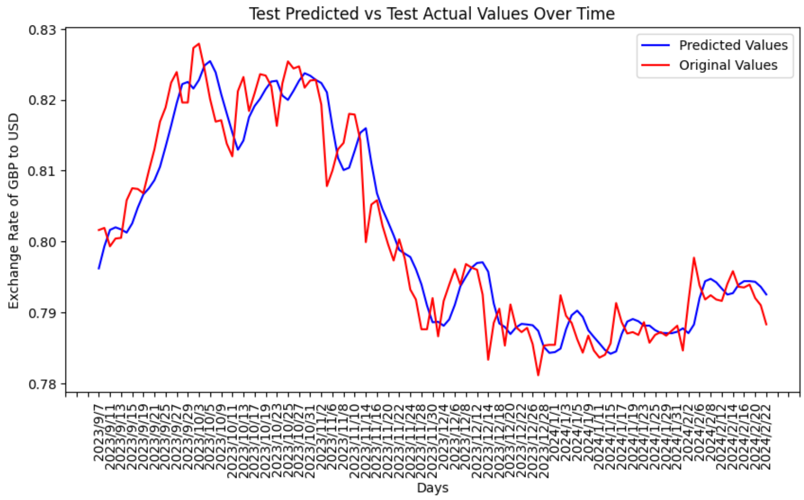


Figure 19. Results of GBP to USD (Picture credit: Original)

LSTM can effectively capture long-term dependencies in time series data through memory unit and gating mechanism, which is suitable for complex time series patterns. It can also adaptively learn the features and patterns of time series data. However, when LSTM deals with large-scale data, the training time is long, the data requirements are also very high, and the predicted data has a certain lag.

3.2. Discussion

This paper selected 3 indexes to evaluate the three models: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R-squared. The comparative results of the indexes of the three models are shown in Table 2 and Table 3.

Table 2. Evaluation of Three Models (JPY to USD)

Model / Index	RMSE	MAE	R-squared
ARIMA	0.2823	0.1457	0.3469
Linear Regression	7.3439	6.3274	0.7669
LSTM	1.0786	0.8805	0.8390

Table 3. Evaluation of Three Models (GBP to USD)

Model / Index	RMSE	MAE	R-squared
ARIMA	0.0017	0.0012	0.5494
Linear Regression	0.0383	0.0312	0.5777
LSTM	0.0041	0.0031	0.9163

In order to compare the prediction accuracy and error variance of the models, the RMSE and MAE values of the three models were observed. Among them, the RMSE and MAE values of ARIMA are the smallest, from which it can be seen that the prediction error of ARIMA is the smallest and the prediction effect is more accurate. In this respect, the predictive ability of ARIMA is higher than that of Linear Regression and LSTM.

In order to compare the degree of fit of the data, the research observed the R-squared values of the three models. The R-squared value of LSTM is closest to 1. It can be seen that LSTM has the highest degree of fitting to the data. In this respect, the data fitting capability of LSTM is better than that of ARIMA and Linear Regression.

4. Conclusion

This research used comparative models to make predictions on foreign exchange rates. The models are ARIMA, Linear Regression, and LSTM. The evaluation indexes chose are RMSE, MAE, R-squared. At last, this research makes a conclusion that ARIMA predicts more accurately than other models while LSTM has the highest degree of fitting to data. The limitation is that this research did not take seasonal and other factors into consideration during the predictions. Although this research mainly focuses on the role of ARIMA, Linear Regression, LSTM in predicting foreign exchange rates, their application potential in other fields is still broad. In the future, the future study will seek to expand the research to explore the potential applications of statistical models and machine learning in other fields such as healthcare, education, and more. This will help deepen the understanding of how statistics and machine learning work in different contexts and provide a theoretical and practical basis for their further application.

References

[1] Bangyuan Z. Critical Comparisons on Deep Learning Approaches for Foreign Exchange Rate Prediction. Retrieved February 29, 2024, from https://arxiv.org/ftp/arxiv/papers/2307/2307.06600.pdf.

[2] Lal M, Kumar S, Pandey DK, Rai VK, Lim WM. Exchange rate volatility and international trade. Journal of Business Research, 2023, 167: 114156.

[3] Qiu Y, Wang J, Jin Z, et al. Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control, 2022, 72: 103323.

[4] Biju A K V N, Thomas A S, Thasneem J. Examining the research taxonomy of artificial intelligence, deep learning & machine learning in the financial sphere—a bibliometric analysis. Quality & Quantity, 2024, 58 (1): 849-878.

[5] Qiu Y, Chang C S, Yan J L, et al. Semantic segmentation of intracranial hemorrhages in head CT scans. 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2019: 112-115.

[6] Colombo E, Pelagatti M. Statistical learning and exchange rate forecasting. International Journal of Forecasting, 2020.

[7] Dautel AJ, Härdle WK, Lessmann S, Seow H-V. Forex exchange rate forecasting using deep recurrent neural networks. Digital Finance, 2020.

[8] Aggarwal P, Sahani AK. Comparison of Neural Networks for Foreign Exchange Rate Prediction. IEEE Xplore, November 1, 2020.

[9] Investing. Today's exchange rate of foreign exchange of foreign currency to real-time query _ the currency market and movements of affection Investing.com Retrieved March 16, 2024, from https://cn.investing.com/currencies.

[10] Hayes A. Autoregressive Integrated Moving Average (ARIMA), December 18, 2022. Investopedia. https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp.

[11] Zach. Introduction to Simple Linear Regression, December 11, 2018. Statology.

[12] Brownlee J. A Gentle Introduction to Long Short-Term Memory Networks by the Experts, July 19, 2017. Machine Learning Mastery.

[13] Chang M, Bae S, Cha G, Yoo J. Aggregated Electric Vehicle Fast-Charging Power Demand Analysis and Forecast Based on LSTM Neural Network. Sustainability, 2021, 13 (24): 13783.