technische universität
dortmund

Quantitative Finance

Term Paper – Summer 2023

Mission: Forecasting Exchange rate

Ly Le Thi; 236938

September 9, 2023

Lecturers: Prof. Dr. Peter Posch

Dr. Daniel Neukirchen

Talble of contents

List of Figure

List of Table

1. Introduction

Exchange rates is the rate at which one currency can be traded for other currency (Krugman, 1979) . When the exchange rate is favorable, exports become cheaper, and imports become more expensive, boosting the competitiveness of domestic producers. Conversely, when the exchange rate is unfavorable, exports become more expensive, and imports become cheaper, hurting the competitiveness of domestic producers. Moreover, exchange rates have a significant impact on investments, particularly those involving foreign currencies. Therefore, exchange rate forecasting is essential for businesses to plan their production and pricing strategies and for investors to make informed decisions and manage risks. Thus, accurate exchange rate forecasting is crucial for businesses, investors, and policymakers to make informed decisions and manage risks in the foreign exchange market, which is the largest financial market in the world with an average daily turnover of $6.6 trillion in April 2019 (Bank for International Settlements, 2023).

In recent years, machine learning and deep learning techniques have gained popularity for exchange rate forecasting due to their ability to capture complex patterns and relationships in the data as well as the increased computational performance. In this term paper, two machine learning methods (Random Forest and eXtreme gradient boosting (Xgboost)) and one deep learning method (Long-Short term memory) are employed to forecast the exchange rates of the US Dollar per Euro, US Dollar per Canadian Dollar and US Dollar per British Pound Sterling based on the exchange rates time series and other factors that affect the movement of exchange rates such as stock data, interest rates (or bank policy rates), consumer price index, economic policy uncertainty, money supply…. The remainder of this seminar paper is structured as follows: In Part 2 the data and its aggregation are described, the methods implemented in the project will be discussed in Part 3 and the results of supervised machine learning methods and deep learning method will be presented in Part 4 respectively. In Part 5, we will compare the performance of all employed models and give a discussion and conclusion of the results.

2. Data

The final data set is based on a 35-time series with different lengths and frequencies and its lags. In order to obtain a feasible data set that also resembles the reality the variables are transformed, and the time series are trimmed. In this section, the base time series will be described before moving on to the description of the transformations and the final data set. Most of the data is provided by the Federal Reserve Bank of St. Louis (FRED) (2023) (for most of the data) and from Investing.com (2023) (for stock data) and based on Baker et al. (2016) for Economics policy uncertainty.

2.1. Main dataset

The primary data employed in this study comprise three daily exchange rate time series. Using the notations of USD: US dollar; CAD: Canadian dollar; GBP: Great Britain pound; and EUR: Euro, then the three exchange rates are 1) USD per CAD, 2) USD per GBP, 3) USD per EUR.

In addition, our full sample period is from April 02, 2018, through to March 31, 2023 (1824 data points), and we split the full period into the training data before January 01, 2022 (about 75% for traning data) and testing data from and after January 01, 2022 (25% for testing data). Figure 1 exhibit the time-series dataset of US dollar per Canadian dollar, per Great Britain pound, and per Euro, respectively and Figure 2 represents the splitting point into tranning and testing set. Moreover, Table 1 provides their descriptive statistics also for the full period.

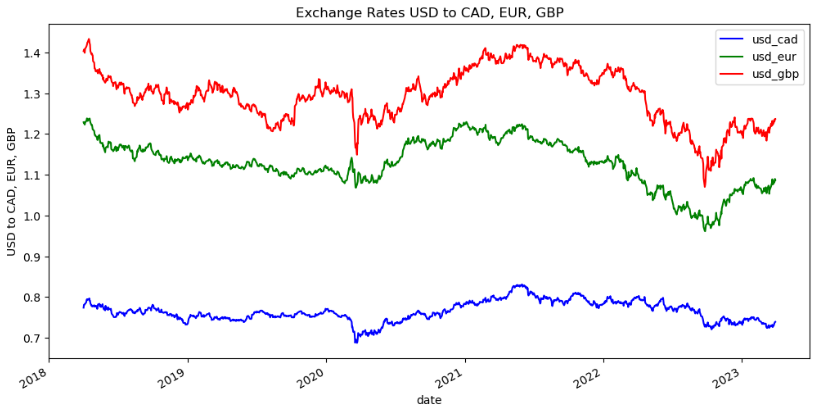

Figure 1. Exchange rates USD per CAD, EUR and GBP

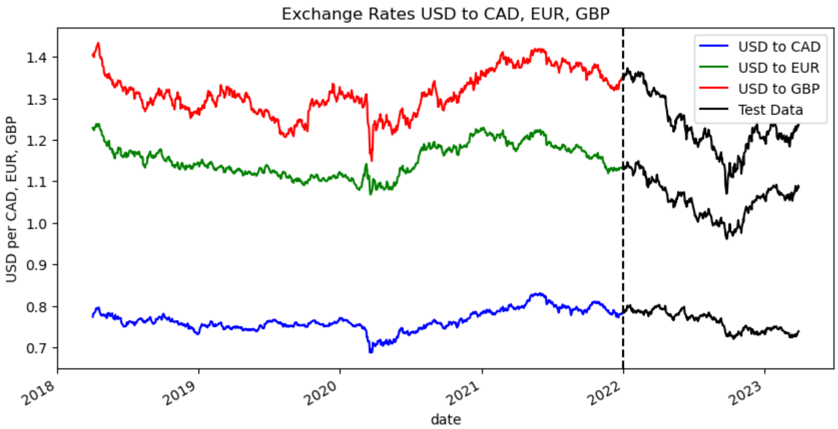

Figure 2. Split training/testing set

Table 1. Descriptive statistic for the full period

	USD per CAD	USD per EUR	USD per GBP
Mean	0.765193	1.127909	1.293042
Standard Deviation	0.025135	0.057725	0.067342
Minimum	0.687805	0.9616	1.0703
Maximum	0.831186	1.2384	1.433200
Skewness	0.052061	0.052061	-0.329925
Kurtosis	-0.056533	-0.056533	-0.228753

2.2. Fundamental factors

The purpose of the project is forecasting three exchange rates with the effect of other factors:

Exchange rate time series with the daily exchange rate data which taken the lag from 1 up to 23 days (1,2,3,4,5,6,7,8,9,10, 19, 20, 21, 22, 23 days lag). These lagged exchange rates serve as crucial time series predictors. They allow the model to capture both weekly and monthly patterns in exchange rate movements. By including lags, models can learn how past exchange rate values influence future rates.

Besides the time series of three exchange rates, the stock data which representative for the economic status of four countries are also included which are the S&P 500 Index (for the U.S.), the S&P TSX Composite Index (for Canada), the STOXX Europe 600 (for Eurozone) and the FTSE 100 (for the U.K.). All these time series are given on a daily frequency, excluding weekends and national holidays. For daily stock data, I took the lags up to 10 trading days.

The project also additionally included variables that capture the macroeconomic indexes and monetary policy in the economic health under consideration. Basic macroeconomic theory would imply that a relative change between interest rates and/or the money supply might lead to depreciation or appreciation of an exchange rate (Venables, 1990). For this purpose, we included monthly data on the policy rates (or interest rates) set by central banks (daily frequency) and the money supply aggregates M1 and M2 (monthly frequency). Besides capturing the monetary policy directly, changing relationships between these variables might indicate a change in the overall pattern of the economy as this would mean that the effectiveness of the measures of the central bank changes. Finally, we included monthly CPI, Unemployment, and quarterly GDP data as indicators for the overall state of the economy of the respective country. Additionally, the Economic Policy Uncertainty Index also included, that is provided on a monthly basis for the 4 countries, to see how uncertainty in policies like trading policies and fiscal decisions can introduce volatility in currency values as traders assess potential economic impacts.

3. Methodology

3.1. Introduction to Supervised Machine Learning

Machine learning is a multifaceted discipline that spans across various domains such as information technology, statistics, probability, artificial intelligence, psychology, and neurobiology. Its core concept revolves around solving problems by constructing models that effectively represent specific datasets by developing algorithms that enable computers to learn from data (Verdhan, 2020). Over time, machine learning has evolved significantly to become a comprehensive field that contributes fundamental computational theories to the realm of statistics.

Supervised learning is widely recognized as the foremost application of Machine Learning. One common technique in this approach involves applying statistical algorithms with historical data to understand and capture the patterns (Verdhan, 2020). This procedure is known as algorithm training. The historical data, or training data, comprises both input and output variables, constituting a collection of training examples that the algorithm learns from. Throughout the training phase, the algorithm establishes a connection between the output variable and the input variables, aiming to formulate a mathematical equation capable of predicting the output variable based on the input variables. Output variables are also referred to as target or dependent variables, while input variables are termed independent variables.

3.2. Supervised machine learning method using in the project.

In recent years, machine learning and deep learning techniques have gained popularity for exchange rate forecasting due to their ability to capture complex patterns and relationships in the data as well as the increased computational performance.

3.2.1. *Random Forest*

Random Forest is a machine learning ensemble method that is widely used because of its flexibility, simplicity, and often quality results. It is a supervised machine learning algorithm that uses multiple decision trees in aggregate to help make more stable and accurate predictions (Aria et al., 2021). In a Random Forest, the algorithm creates multiple sub-models, each represented by a decision tree, which work independently of each other. Each sub-model randomly selects a subset of the data with replacement and a subset of the available variables when deciding how to branch nodes. To make a final prediction, the Random Forest algorithm aggregates the results from all these independent sub-models. This ensemble approach often results in more accurate and robust predictions compared to a single decision tree, as it combines the knowledge learned from multiple perspectives on the data.

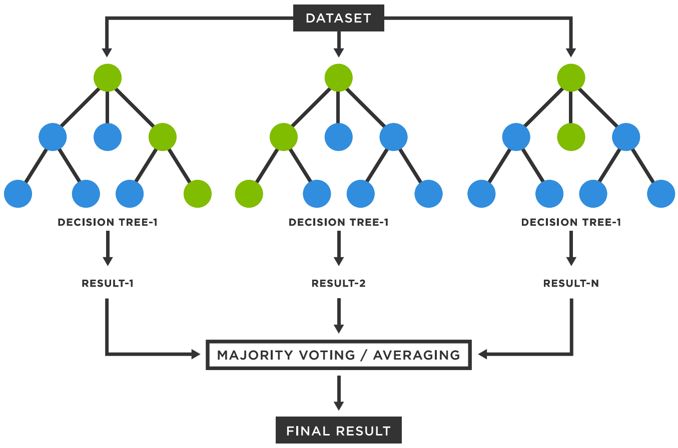

Figure 3. Structure of Random Forest (Tibco.com, 2023)

- **Random Forest for Forecasting exchange rates**

In this project, I use Random Forest regression to build the model for forecasting each type of exchange rate with maximum 5000 trees, maxdeep is 10 in each forest and a fixed random state for reproducibility. To optimize the models, I also perform hyperparameter tuning. Hyperparameter tuning in a Random Forest model aims to find the optimal combination of hyperparameters to create the best-performing model for prediction task. After fitting the models with the training data and corresponding target exchange rate values, the best-performing model result after hyperparameter tuning is used to make exchange rate predictions in testing area. This approach allows the research to leverage the power of Random Forest regression, capture the relationship between exchange rates and various influencing factors, and select optimal hyperparameters to improve the accuracy of the exchange rate forecasts.

3.2.2. *Extreme Gradient Boosting (Xgboost)*

XGboost (eXtreme Gradient Boosting) is a popular and powerful machine learning algorithm that falls under the category of gradient boosting methods (Brownlee, 2016). Boosting is an ensemble learning technique in which new models are added to correct mistakes made by existing models for better predictive performance. Models are added sequentially until no further refinement can be made (Brownlee, 2021). XGboost implements the gradient boosting decision tree algorithm (or shortly gradient boosting) and specifically designed for solving both classification and regression tasks. It is called gradient boosting because it uses a gradient descent algorithm to minimize the loss when adding new models (Brownlee, 2016). Unlike the decision tree, in gradient boosting, making a tree starts with a leaf, which means that each tree model in XGBoost minimizes the residual from its previous tree model. The initial leaf is filled with the average value of the features to be predicted. XGBoost is known for its high performance and efficiency.

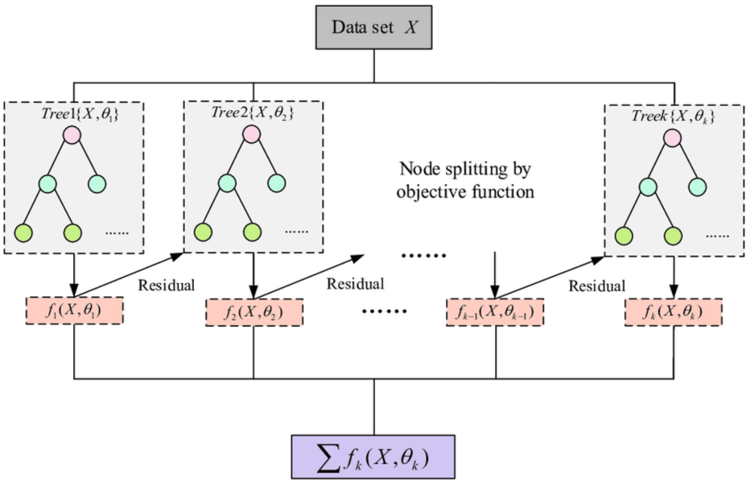

Figure 4. Structure of Xgboost (Guo et al., 2020)

- **Xgboost for Forecasting exchange rates**

First, a base XGboost model is configured with some initial settings (maximum 10000 decision trees, maxdeep is 10). The building the model process is optimized by applying hyperparameter tuning. Hyperparameter tuning in XGBoost works by systematically searching for the best combination of hyperparameters to improve the performance of the model. The goal is to find the set of hyperparameters that results in the best predictive performance on the testing dataset. Then, fit the XGBoost models separately to the training data for each exchange rate of interest. This step allows the models to learn the relationships between historical data and other influencing factors. The best model will be used for prediction on Testing data of each exchange rate.

3.3. Deep learning - Long Short-Term Memory

Long Short-Term Memory (LSTM) is a specialized type of recurrent neural network designed for sequential data and time series analysis. LSTMs are widely used in natural language processing, speech recognition, and time series forecasting, particularly dealing with long-range dependencies. They excel at capturing complex patterns and retaining information over extended sequences, making them suitable for language modeling, translation, and speech recognition (Brownlee, 2017).

LSTMs feature memory cells capable of storing and processing data over long sequences, mitigating the vanishing gradient problem found in basic RNNs (Smagulova & James, 2020). They employ gates, including input, forget, and output gates, to control data flow within the network. Input Gate decides how valuable the current input is to solve the task. Forget Gate decides which current and previous information is kept and which is thrown out and in the Output Gate, the output of the LSTM model is then calculated in the Hidden State. The structure of LSTM is present in Figure 5.

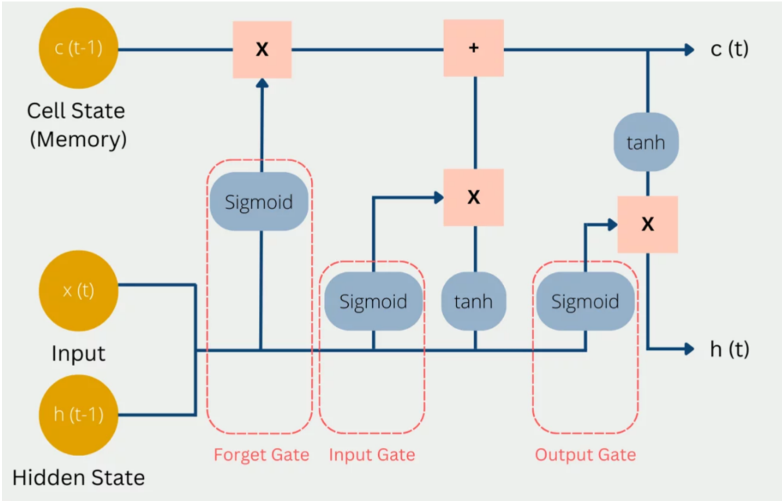

Figure 5. Structure of Long-Short Term Memory (Brownlee, 2017)

- **LSTM for Forecasting exchange rates**

This project applies LSTM by designing a flexible LSTM model with tunable hyperparameters for each target exchange rate (USD per CAD, USD per EUR, USD per GBP). The training data then be devide 70% for training into and 30% for validation. I optimize hyperparameters by changing LSTM units and learning rates to minimize mean squared error. The model then will be used to train the data in the training set to capture the complex relationships between lagged exchange rates and various influencing factors and used it for prediction. The process is quite time-consuming, especially with large trials. After having well-tuned LSTM models, the predictions are made accurately on the testing dataset, providing valuable insights into future exchang rates movements.

3.4. Evaluation

After forecasting for all models, I evaluate the predictive performance of the random forest, Xgboost, and LSTMs approach by the R-square and Roots Mean Square error (RMSE), which is a very reasonable measure for evaluating the predictabilities of exchange rate levels investigated in this study. RMSE is the square root of the mean of the squared residuals, it shows the "average distance" of the observations from the line of best fit and can be calculated as follows: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \varepsilon^2}$ where n denotes the total number of observations, and *ith* is the residual ε for each observation. A high RMSE implies that the fitted values are further away from the actual data.

The R^2 is a measure which is used to assess to what extent the variation of the response variable can be explained by the model: (R^2 = explained variation / total variation). If the R^2 is too small, the model might be incomplete, some important variables are missing. Usually, a large R^2 is desirable, however, it can mean that there is overfitting in the model.

To establish the evaluation of the model's performance, this research also constructs a baseline model, where the value at time (t-1) is used as the predictive value for time t ($Y_t = Y_{t-1}$). Consequently, the baseline model demonstrates a proficient capability in forecasting exchange rate values. The R-square value of predicted value against the testing value approaches 1 for all exchange rates, indicating a high level of explanatory power, while Root Mean Squared Error (RMSE) exhibit minimal magnitudes (Table 3), further corroborating the model's effectiveness.

4. Forecasting Results

4.1. Hyperparameters

Hyperparameter tuning is the process of finding good model hyperparameters, that reduce the RMSE. This value search takes a longer time than running the core model. The hyperparameter values looking for are n_estimators, n_jobs, max_depth, x_colsample_bytree, x_gamma, x_min_child, x_reg_lamda, and x_subsample. For random forest and Xgboost, searching for hyperparameter is supported by *RandomizedSearchCV()* function in Python with 3-fold cross validation and for LSTM the searching is optimized by using *RandomSearch()* with max_trials is 100 to maximize the chance to find the best hyperparameters . After a successful search, results for each model are presented in Table 2.

Table 2. Hyperparameter results

Random Forest	**Optimal Hyperparameter**
Model for CAD	`{'n_estimators': 5000, 'max_features': 'auto', 'max_depth': 10}`
Model for EUR	`{'n_estimators': 3666, 'max_features': 'auto', 'max_depth': 20}`
Model for GBP	`{'n_estimators': 2777, 'max_features': 'auto', 'max_depth': 20}`
Xgboost	
Model for CAD	`{'n_estimators': 2000, 'max_depth': 1, 'learning_rate': 0.01}`
Model for EUR	`{'n_estimators': 4000, 'max_depth': 1, 'learning_rate': 0.05}`
Model for GBP	`{'n_estimators': 2000, 'max_depth': 1, 'learning_rate': 0.01}`
LSTM	
Model for CAD	`Optimal number of units in the first LSTM layer is 192` `Optimal learning rate for the optimizer is 0.005063713408586714.`
Model for EUR	`Optimal number of units in the first LSTM layer is 480` `Optimal learning rate for the optimizer is 0.0012281783364939784.`
Model for GBP	`Optimal number of units in the first LSTM layer is 384` `Optimal learning rate for the optimizer is 0.004325833653269832.`

4.2. General results on testing data

Table 3. Result for R^2 and RMSE on testing data

		USD per CAD	**USD per EUR**	**USD per GBP**
R^2	**Baseline model**	0.9750	0.9796	0.9831
	Random Forest	0.9705	-0.3406	0.6825
	Xgboost	0.9641	-0.2034	0.7796
	LSTM	-1.3864	-0.5035	0.5529

RMSE	**Baseline model**	0.00377	0.29685	0.47303
	Random Forest	0.00410	0.05259	0.03864
	Xgboost	0.00452	0.32975	0.48621
	LSTM	0.03689	0.34138	0.49358

In the table above, we can see the R^2 and RMSE values on testing data for the three different model for three different currencies pair. From the 3 models, only random forest model and Xgboost can achieve similar performance as the baseline model and that only for the USD-CAD exchange rate. LSTM perform poorly for USD-CAD. Three models have average fitting for USD-GBP and quite poorly performed for USD-EUR The results for RMSE from the predicted models show that the results are again quite good for USD -CAD, when all three models resulting in a very small RMSE, similar performance as the baseline model. For USD-EUR and USD-GBP, the results of RMSE for random forest is even smaller than baseline model but the results for Xgboost and LSTM is higher.

4.2.1. *Feature importance*

Feature importance is also calculated for each currency (CAD, EUR, and GBP) in Random Forest and Xgboost and it's based on the Gini importance score. Gini importance is a measure of how much each feature contributes to the decision-making process of model. The importance scores are normalized such that the most important feature has a score of 1.0, and lower importance features have scores relative to the most important feature. Assessing feature importance in each model for forecasting various exchange rates has revealed key contributing factors. Among these factors, it is evident that the historical data of the exchange rate itself (lag 1, 2, 3 days) has the most substantial impact on the predictive models. For example, USD-CAD lags 1 day contribute 0.909368 in the importance follow by the USD-CAD lags 2 and 3 days. Additionally, the stock indices representing both countries involved in the exchange, play a significant role in the overall model. The historical values of competitor exchange rates, as well as Consumer Price Index, Unemployment rates, and the Economics policy uncertainty also exhibit notable contributions to the model's performance.

4.3. Forecating visualization

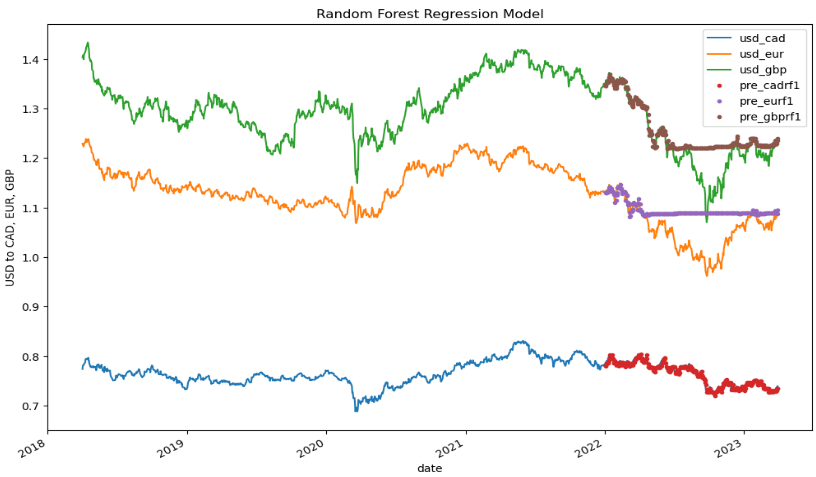

Figure 6. Predicted value from Random Forest model.

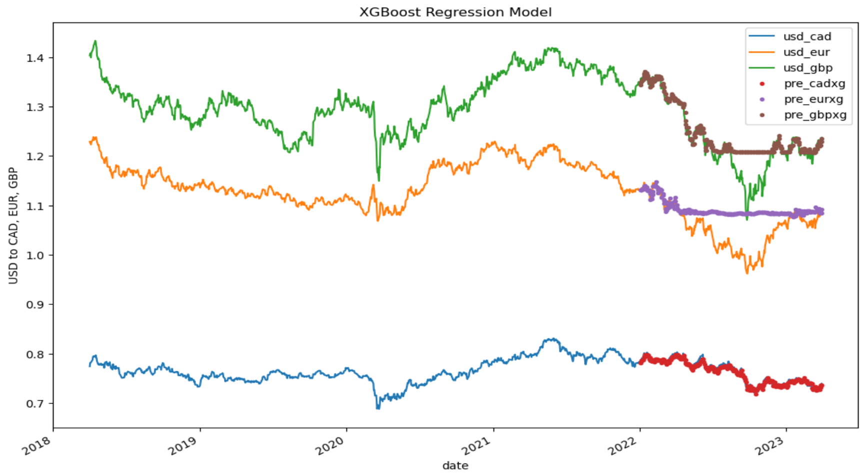

Figure 7. Predicted value from Xgboost model.

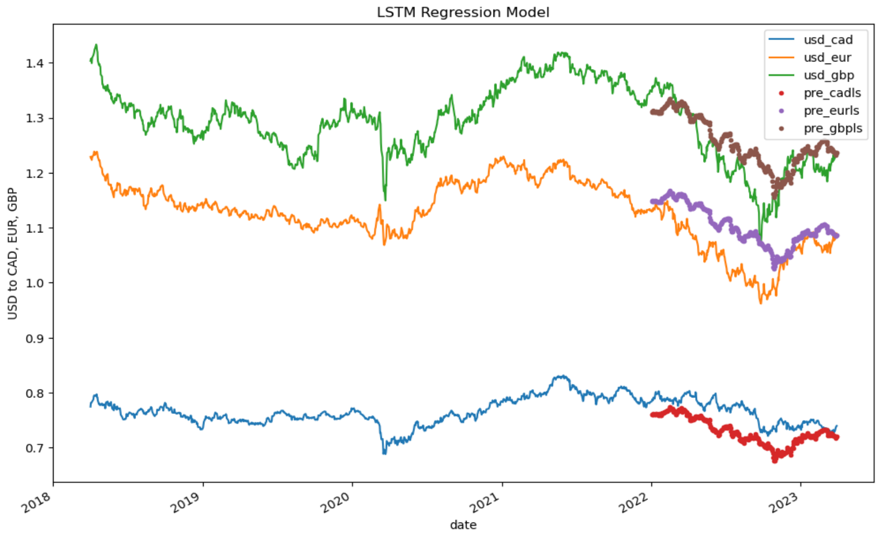

Figure 8. Predicted value from LSTM model.

Figure 6, Figure 7 and Figure 8 depict the predicted values, generated by the three models: Random Forests, XGBoost, and LSTM, for different types of exchange rates, against the real value. In the case of the Random Forest model, the predictions exhibit a high degree of accuracy for the CAD exchange rate, an intermediate level of fit for GBP, but less satisfactory performance for EUR. A similar pattern emerges in the results of the XGBoost model, which demonstrates high precision for the CAD exchange rate, average fitting for GBP, but less favorable outcomes for EUR.

Additionally, it's worth highlighting that the results for the CAD exchange rate underscore the effectiveness of Random Forest and XGBoost models in handling stable or not so much fluctuating data. Consequently, the predictions for CAD exhibit a high level of precision, with minimal discrepancies between R-square and RMSE values between the training and testing datasets. The outcomes derived from the Long Short-Term Memory (LSTM) model and the associated visualizations reveal that LSTM effectively captures the extended temporal dependencies within the time series, considering the influence of various factors and lag effects within the time series. The patterns identified by LSTM are notably accurate, although it's important to note that the overall fit of the time series is not consistently strong across all three models.

4.4. Limitation of 3 models in forecasting

4.4.1. *Overfitting-Underfitting*

Xgboost and random forest model experienced overfitting due to their strong performance on the training data compared to those in testing data depending on R^2 and RMSE results for both training and testing data. For example, the Random Forest model displayed high R-squared values in training data for USD per CAD, EUR, and GBP (0.9973, 0.9982, and 0.9975, respectively), surpassing baseline and testing data results (except the result for USD-CAD which the result is close). A similar pattern emerged with XGBoost and LSTM models, exhibiting higher R-squared values and smaller RMSE in training data than in testing data. Efforts to alleviate overfitting in the Random Forest and Xgboost models by reducing tree sizes or excluding variables led to reduced model performance compared to the original setup, taking cross-validation in choosing hyperparameters using *RandomizedsearchCV().* An attempt to use cross-validation from splitting data is made but the results are not so good for time series to do random cross-validation. LSTM's less-than-anticipated performance may be attributed to its demanding data requirements for effective fitting, alongside the time-intensive nature of model fitting and hyperparameter optimization.

4.4.2. *Data limitation*

A noteworthy observation from the fitting analysis is that these models struggle to learn and predict the sharp decline in exchange rates that occurred from mid-2022 to the end of 2022, particularly in the EUR and GBP exchange rates. This significant drop was either absent or less pronounced in the training data, explaining the models' suboptimal performance during this period. To address this, an attempt was made to modify the prediction time by shifting the splitting point to October 2022, coinciding with the point at which the exchange rates began to recover. This adjustment yielded notably accurate results for all exchange rates. In this research I tried to find the data in the 2nd and 3rd quarter for GDP, CPI and Interest rates in 2023 to extend the training data to have reasonable splitting point but the data for 2nd and 3rd quarter has not published yet.

5. Discussion and Conclusion

In this study, an examination of the predictive capabilities of three distinct models—Random Forest, XGBoost, and LSTM—was conducted for three different exchange rates: CAD, EUR, and GBP. The evaluation was based on performance metrics encompassing R^2, RMSE, and fitting visualization.

In general, XGBoost appeared to outperform the other two models (Random Forest and LSTM) in forecasting exchange rates, especially for USD-CAD. The EUR exchange rate exhibited greater volatility over time and experienced a pronounced and extended decline during the testing period (mid-2022 to the end of 2022). This substantial variation posed a challenge for predictive models, as they had not adequately captured this specific drop in their training data, leading to diminished predictive performance. In the case of the GBP exchange rate, all three models displayed a range of performance from average to good. LSTM, while proficient at learning the intricate patterns within exchange rate data, exhibited a limitation in achieving precise fitting and forecasting. This observation suggests that LSTM may excel in capturing the temporal dependencies but may struggle with fine-tuned prediction.

In practice, the results from 3 models show that all the machine learning models we explored in the report perform worse than a very simple baseline model: Apparently, today's exchange rate is the best predictor of tomorrow's exchange rate, outperforming all the complex approaches. This is not a surprise: If anyone could predict exchange rates better than the current exchange rate can, that would suggest that there are major inefficiencies in the global exchange rate marketplaces. This is likely not the case because anyone who has such a model that can predict future exchange rates can become a multi-millionaire in no time, by buying and selling currency. However, exploring and improving the forecasting model is needed to find the optimal model for futher forecasting. For the future research, I will explore enhanced cross-validation techniques suitable for time series data, as well as the utilization of more extensive datasets to bolster the LSTM model's predictive capabilities.

6. References

Abdulhafedh, A. (2022). Comparison between common statistical modeling techniques used in research, including: Discriminant analysis vs logistic regression, ridge regression vs LASSO, and decision tree vs random forest. *Open Access Library Journal, 9*(2), 1-19.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics, 131*(4), 1593-1636.

Bank for International Settlements. (2023). Foreign exchange turnover in April 2019. Retrieved 01/09/2023, from https://www.bis.org/statistics/rpfx19_fx.htm

Brownlee, J. (2016). *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*: Machine Learning Mastery.

Brownlee, J. (2017). *Long short-term memory networks with python: develop sequence prediction models with deep learning*: Machine Learning Mastery.

Brownlee, J. (2021). *Ensemble learning algorithms with Python: Make better predictions with bagging, boosting, and stacking*: Machine Learning Mastery.

Federal Reserve Bank of St. Louis (FRED). (2023). Federal Reserve Economic Data | FRED | St. Louis Fed. Retrieved 20.08.2023, from https://fred.stlouisfed.org/

Guo, R., Zhao, Z., Wang, T., Liu, G., Zhao, J., & Gao, D. (2020). Degradation state recognition of piston pump based on ICEEMDAN and XGBoost. *Applied Sciences, 10*(18), 6593.

Investing.com. (2023). Stock Index. Retrieved 20.08.2023, from https://de.investing.com/indices/stoxx-600-historical-data

Krugman, P. R. (1979). Vehicle currencies and the structure of international exchange: National Bureau of Economic Research Cambridge, Mass., USA.

Smagulova, K., & James, A. P. (2020). Overview of long short-term memory neural networks. *Deep Learning Classifiers with Memristive Networks: Theory and Applications*, 139-153.

Tibco.com. (2023). Retrieved 05.09.2023, from https://www.tibco.com/reference-center/what-is-a-random-forest

Venables, A. J. (1990). Microeconomic implications of exchange rate variations. *Oxford Review of Economic Policy, 6*(3), 18-27.

Verdhan, V. (2020). Supervised learning with python. *Okänd. Irland: Apress*.