Forecasting Exchange Rates Using Machine Learning Methods

LAU Yuen-Hoi, YU Ho-Nam
MAU Yuk-Hung, YIU Wai-Keung Binnie
Hong Kong University of Science and Technology

Abstract

It is no doubt that foreign exchange rates play an important role around the world. Accurate predictions of exchange rates serve as valuable reference when not only countries but also you and me to make investment decisions. In this connection, we studied the feasibility of application of machine learning methods in exchange rate forecasting (Euro/USD) in this project. Thanks to the transparency of financial figures nowadays, we first collected all related financial data from the internet. Second, we made use of the collected data to review the forecasting performance of various popular machine learning methods such as OLS, SVM, Random Forest, XGBoost, LSTM and temporal CNN. Third, we proposed to add attention layers to temporal CNN to improve its performance. In this project paper, we illustrated the results of different methods in terms of prediction errors including MAE and RMSE, prediction accuracy rates of change of direction, and cumulative profits generated from these models. Among all the methods we tested, linear regressions perform the best. We conclude that machine learning methods have varying forecasting capability regarding exchange rates.

1. Introduction

Forecasting exchange rates has important implications for the government and investors. From the government's point of view, keeping the volatility of exchange rates to a low extent is crucial to the economy. If the home currency appreciates to a large extent, there is a large amount of capital inflows, which may influence the asset prices greatly. On the contrary, when the home currency depreciates tremendously, huge amount of capitals flow out of the country, leading to a negative impact on the economy. Thus, the government can come up with some measures to combat the huge volatility of exchange rates before it happens. From investors' point of view, they can make a fortune if exchange rates can be forecast accurately. In addition, investors can avoid great depreciation of the currencies they are holding if they can sell those currencies before it happens.

There are a large volume of literature regarding forecasting exchange rates using linear or non-linear regression methods with economic variables such as interest rates and inflation rates as predictors. However, most of the findings are that those methods cannot outperform the random walk model without drift in terms of root mean squared errors (RMSE) and mean absolute errors (MAE). In recent years, thanks to the rapid developments of machine learning models, some papers tried to forecast exchange rates via machine learning methods. However, the findings are mixed. While there are related works about applying shallow machine learning models in forecasting exchange rates, there is only a few literature about exploring deep learning models on this area.

In view of this, this project can make contributions to the literature by exploring both shallow machine learning models and deep learning ones in forecasting exchange rates. We have two forecasting tasks in order to exam the predictability of exchange rates using machine learning methods including OLS, random forest, XGBoost, SVM, temporal CNN and our proposed temporal CNN with attention. First, we aim to forecast the monthly spot values such as Euro in terms of US dollars (EUR/USD). Second, we aim to predict the direction of change of the value of currency pairs in future periods. Furthermore, we construct a virtual investment portfolio using historical data according to the predicted direction of change in future periods and examine the cumulative returns of the portfolio constructed from forecasting methods. We gather the data from September 2003 to February 2021 and we used the first half of the data as the training set and the remaining half as the test set. Details of the data will be described in the data section.

This project report is structured as follows. In section 2, the problem of forecasting exchange rates is defined. Section 3 shows related published works that are relevant to our project and how our approach differs from existing works. In section 4, machine learning models are introduced . Section 5 describes the dataset of exchange rates and relevant economic variables. Section 6 shows the experiment results of our model. Section 7 concludes our project paper.

2. Problem Statement

The problem we are trying to solve is to forecast the exchange rates of a currency pair in future periods as accurately as possible.

We have two forecasting tasks. The first one is formulated through this equation: $y_{t+1} = f(X_t)$. We would like to forecast the exchange rates in the next period denoted as y_{t+1} using all relevant historical data at present period denoted as X_t. The function f is a black box that we would like to explore. In our project, we use machine learning models including deep learning ones to forecast the next-period exchange rates. Evaluation metrics include MAE, RMSE and R-Squared.

The second forecasting task is to determine the change of direction (i.e. up or down) in the next period.

$$up_{t+1} = \begin{cases} 1 \ \ if \ \ \hat{y}_{t+1} \geq y_t \\ 0 \ , \ \ \text{otherwise.} \end{cases} \tag{1}$$

The exchange rates are predicted to go up if the forecast values are greater than or equal to the current values. Otherwise, "down" direction is predicted.

3. Related Work

To address the exchange rates forecasting problem, there are existing traditional models utilizing the historical information of economic variables such as interest rates, inflation rates, GDP growth and commodity prices to predict future exchange rates. Ordinary Least Squares (OLS), Autoregressive Integrated Moving Average (ARIMA), GARCH are among those models. However, they can only capture shallow features. Univariate deep learning models including Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) models can capture the deep temporal features in single time series only. In this paper, we would like to explore the predictability of not only traditional machine learning models, but also newly developed ones. Temporal Convolutional Neural Network (TCNN) can capture deep characteristics of multivariate time series data.

In the economic stream of predicting exchange rates, [12], [10], [20], [4] and [16] explore the OLS using economic variables. Most of the prediction results are that models cannot outperform random walk without drift in next-period values prediction. [20] includes more quarterly data such as Taylor rule fundamentals, yield curve factors, shadow rates and risk and liquidity factors in the OLS models and find that they can outperform the random walk in the direction-of-change predictions. Similar conclusions are found in [11] while using weekly data to make predictions. It is hard to beat random walk in terms of next period value prediction. However, the OLS model performance greatly improves in the direction-of-change metric. [16] answers the question of exchange rate predictability by saying "it depends". A linear model with a small number of parameters and Taylor rule factors as the predictors performs better than alternatives. [5] built factor models to forecast quarterly exchange rates and concluded that they can outperform random walk in long horizons of 8 to 12 quarters. [13] examined the predictability of EUR/USD using Artificial Neural Network (ANN), ARCH and GARCH. The paper found that ARCH and GARCH are better than ANN when predicting the dynamics of EUR/USD.

In the area of applying machine learning methods to forecast exchange rates, [15] examined SVM, ANN and GBC whether they can generate profitable foreign exchange long/short investment portfolio. The results are that almost all models can make profits. In addition, [3] concludes that when economic fundamentals are inputs to machine learning models, forecasting accuracy rates improve for short term. [17] found that when adding commodity prices as a predictor to random forest, SVN and ANN, random forest obtains a satisfactory forecasting capability and outperforms alternative models.

In this project report, we use OLS, random forest, XGBoost, SVM, LSTM, Temporal CNN(TCNN) and our proposed TCNN with attention to predict the next-period EUR/USD. Fifty percent of the dataset is employed as the ground truth to quantitatively evaluate our methods. Common evaluation metrics such as Mean Absolute Errors(MAE), Root Mean Squared Errors(RMSE) and R-Squared are used in our case.

4. Models

- eXtreme Gradient Boosting (XGBoost) [6]: Boosting algorithm improves the prediction by training a sequence of weak models. Trees are built sequentially by predicting the difference of the actual target value and the predicted value from previous trees. Finally, it can make prediction by weighted sum of the prediction from all the trees . XGBoost is a boosting algorithm. In particular, it considers the regularization term which imposes a penalty on the complexity of the loss function to prevent overfitting. The number of trees is 2000, the maximum depth is 3 and the learning rate is 0.1.

- Elastic Net (EN) [21]: L1 and L2 regularization in EN has been introduced as penalty linear regression which can reduce overfitting . The L2 regularization can produce shrinkage effect on the coefficients which is not important to the predicted target. The features can be eliminated by the sparsity of L1 regularization. Elastic Net combines feature elimination from L1 regularization and feature coefficient reduction from the L2 regularization to improve model's predictions. The L1 ratio is 0.5.

- Support Vector Machines (SVM) : it is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. It is based on statistical learning theory and was developed by Vapnik in the year 1995. The primary aim of this technique is to project nonlinear separable samples onto another higher dimensional space by using different types of kernel functions. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.

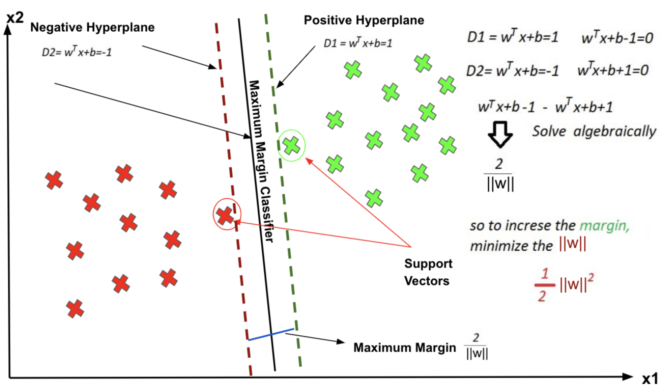

Figure 1. SVM

- Random Forest Regressor (RFR): RFR is one of the basic regression models firstly introduced in 1995 adopted by the trend to predict exchange rate.[8] [14] In RFR, the number of trees in the forest is one of the parameters for adjusting. Theoretically, the higher the value set, the better performance obtained. We used the default value (i.e. 100) in this project because the higher value like 200 and 300 were tested, but, the improvement is insignificant while the process time increased obviously. Default value of other parameters were also adopted.

- Ordinary Least Squares: it is a classic linear regression model. The next period exchange rate (y_{t+1}) is the dependent variable and all other current variables (X_t) are independent variables in our setting: $y_{t+1} = c + \beta X_t$

- Autoregressive Integrated Moving Average (ARIMA): it is a popular time series model for financial related prediction.[1] There are three parameters under ARIMA namely p meaning the number of lag observations included in the model; d meaning the number of times that the raw observations are differenced, also called the degree of differencing; and q meaning the size of the moving average window, also called the order of moving average. The model in this project, the ARIMA(p, d, q) is set as ARIMA(5, 1, 0) as the Akaike information criterion (AIC) is the lowest.

- Gated Recurrent Unit (GRU): it is one of the family members of recurrent neural networks since 2014.[7] In a nutshell, GRU is a simplified LSTM because GRU only consists of two gates namely reset gate and update gate. The advantages of GRU are easier to modify and faster convergence speed.

- Long Short-Term Memory (LSTM) [9]: LSTM is a type of Recurrent Neural Network. A common LSTM consist number of cells, A cell consist of a forget gate, an input gate and also an output gate Figure 2. The cell can remember values over time intervals. The input determines how much to add into current cell state from the input of current input and previous hidden state. The output gate determines the value of the next hidden state. The forget gate determines how much to remember from components of previous cell state. Finally the output from these 3 gates will combine together and the neural network can decide which kinds of information need to be forget or reinforce. This memory property makes LSTM powerful and able to work with longer sequence data. It is shown that LSTM really work better than simple RNN in many cases.

  In this project, LSTM with a new loss function below has been tried. This loss function includes the residual of the true target of next month and current rate and that of the predicted target of next month and current rate. This function likes an angular similarity function. The Loss is the mixture of L1 Loss and this similarity function. The LSTM with L1 Loss has been implemented (denote as Unsim LSTM). The hidden size is 250, the number of recurrent layers is 3.

$$\text{Similarity} = \frac{(y_{true} - y_{current}) \cdot (y_{pred} - y_{current})}{\|y_{true} - y_{current}\| \, \|y_{pred} - y_{current}\|}$$

- Temporal Convolutional Neural Network (TCNN) [2]: TCNN is a type of neural network which can predict time sequences using convolution structure. To ensure the output with the corresponding dimension for different input in different time step and no data leakage, TCNN consists 1-D Fully-Convolutional Network (FCN) and Dilated Causal Convolutions and it can receive any sizes of sequence data and output a same size of data by using padding. FCN can ensure the output

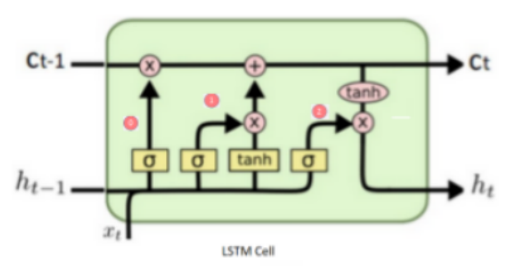

Figure 2. LSTM

with the same number of channels as the input from previous layer. Dilated Causal Convolutions can ensure the output at time step t+1 consider the input at time steps before t+1 only. Residual block is added to prevent vanish gradient problem. The temporal length is 16, the number of hidden units is 12 and the kernel size is 3 for TCNN.

- TCNN with temporal attention added is proposed in this project to improve the forecasting capability of TCNN. The structure is illustrated in Figure 3. The scaled dot-product attention is from [18] which is illustrated in Figure 4. The query is the last hidden unit h_t. We selected previous 16 periods of hidden units as the keys and values. The final output $\tilde{y}_{t+1}$ is calculated as the attention weight times the values. The temporal length is 16, the number of hidden units is 12 and the kernel size is 3 for TCNN with attention.

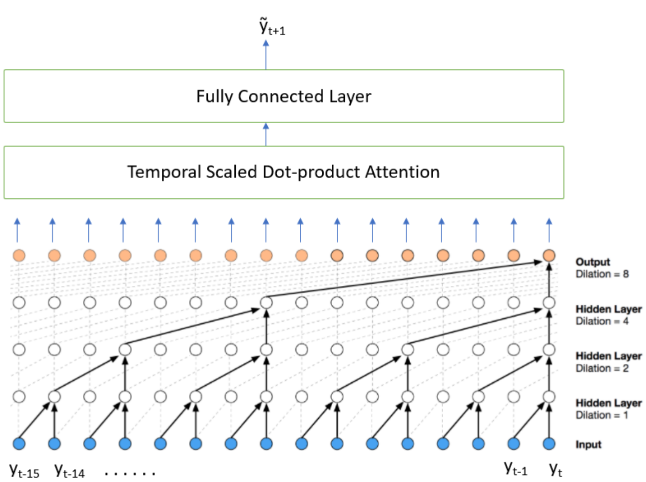

Figure 3. TCNN with temporal attention

5. Datasets

The scope of this project is forecasting the exchange rate between the the EURO (EUR) and the US Dollar (USD), which is EUR/USD. Euro is expressed in terms of USD.

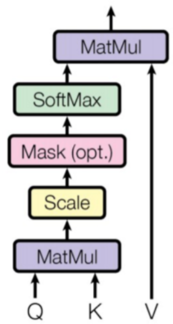

Figure 4. Scaled dot-product attention

Table 1. MARKET FEATURES

Type	Country	Features
Stock Market	US	S&P 500 Index
	Europe	EURO STOXX 50 Index
Asset	Global	Gold Price Brent Oil Futures

When the exchange rates increase, Euro appreciates. This project can be extended to other currencies.
The prediction target is the monthly rate of EUR/USD. We will explore the prediction power through two types of features - Market features and Fundamental features. The market features is the indicator about which has close relationship with traded currency. On the other hand, The Fundamental features is the indicator which reflects the economic status of the corresponding regions. The dataset is collected between September 2003 and February 2021.

A. Market Features
The Market features across both Europe and the United States has been selected. The EURO STOXX 50 Index is the stock index for the Eurozone, which can provides a blue-chip sectors in the region. The S&P 500 Index includes 500 leading companies in US which cover 80% of available market capitalization. The gold is an asset which can consider as inflation hedge. The Crude Oil is an essential component of fuel, plastic product and petroleum products. The oil price can reflect the price of commodities, which can as an indicator of inflation. A summary is shown in Table 1.

B. Fundamental Features
The Fundamental features across both Europe and the United States has been selected. The Interest Rate is the rate of charged amount on the loaned money. The Consumer Price Index (CPI) measures the purchasing power of the currency which can considered as indicator of the inflation or deflation affecting consumers. A summary is shown in Table 2.

Table 2. FUNDAMENTAL FEATURES

Type	Country	Features
Interest Rate	US	Short-term Interest Rate Long-term Interest Rate
	Europe	Short-term Interest Rate Long-term Interest Rate
CPI	US	USA CPI
	Europe	EU CPI

C.Data Processing

In order to fulfill the practical use purpose, the dynamic models have been built instead of static models. The data has been splitted into train set and validation set. The train set includes the period of September 2003 to December 2011 and the additional period up to the previous month of prediction month. The validation set includes the period of January 2012 to February 2021. These subsets are kept in order to maintain the time series nature.

To predict the currency of next month, the features are lagged by one month as the fundamental features are always released lately by a month. For the feature transformation, the difference of Interest Rate has been calculated as a feature since the Interest Rate between two regions usually can be compared directly.

Correlation Heatmap has been computed for the relationship between covarites Figure 5. The current EUR/USD would have the strong correlation with the predicted EUR/USD which is close to 1. The Brent Oil has the second highest correlation with the predicted EUR/USD which is close to 0.8. The Gold Price and CPI have a correlation around 0.5.

Besides the currency prediction, the prediction can be further produced into up and down classification and cumulative return. For up and down classification, a variable can be introduced. It can be set as 1 if the rate of next month is higher the rate of this month and otherwise set as 0.

6. Experiment Results

By the order of time, we use the first 50% data as training set and start to predict next monthly exchange rates recursively until January 2021.

6.1. Experiment Results

6.1.1 Prediction Error Result comparisons for the selected models

The results in Fig. 6 show the overall results of different prediction error metrics for each selected model. We have used random walk model as a reference for other models.

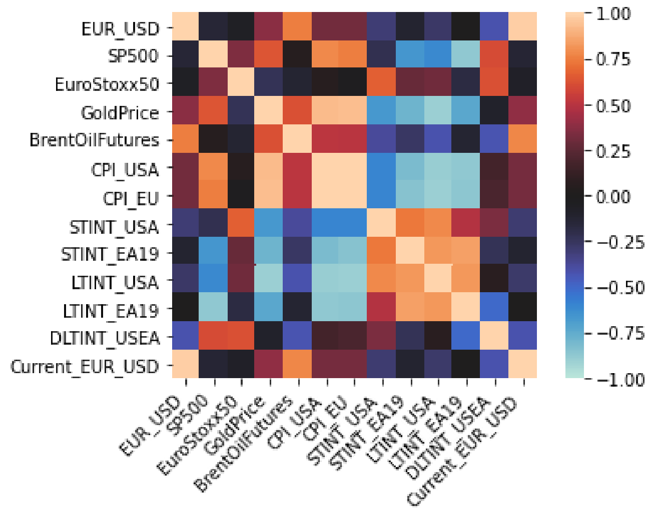

Figure 5. Correlation Heatmap between covariates

	MAE	MSE	RMSE	R-squared
Random Walk	**0.0198**	**0.0007**	**0.0257**	**0.9284**
SVR	0.0681	0.0070	0.0838	0.4943
Random Forrest	0.0232	0.0009	0.0300	0.8962
Linear Regression	0.0235	0.0009	0.0305	0.8974
Unsim LSTM	0.0312	0.0015	0.0389	0.8330
LSTM	0.0369	0.0026	0.0505	0.7758
Xgboost	0.0274	0.0012	0.0347	0.8675
TCNN	0.0621	0.0064	0.0802	0.5256
TCNN with attention	0.0939	0.0129	0.1134	0.2573
Elastic Net	0.0686	0.0071	0.0845	0.3940
ARIMA	0.0194	0.0006	0.0241	0.7143
GRU	0.3113	0.1208	0.3476	0.5182

Figure 6.

Mean Absolute Error (MAE) refers to the absolute difference between our model's predictions and the ground truth, then average it out across the whole dataset. So a smaller MAE means the model performed better in prediction tasks.

Mean Squared Error (MSE) is perhaps the simplest and most common loss function. To calculate the MSE, you take the difference between your model's predictions and the ground truth, square it, and average it out across the whole dataset. A small MAE imply a good prediction model.

Another most commonly used metric for regression tasks is root-mean-square error (RMSE). This is defined as the square root of the average squared distance between the actual score and the predicted score.So it is relative to the variance in target values.

R-squared (R^2) is the coefficient of determination regression score function. Predicted values are used instead of the fitted values in the original R^2 calculation. R^2 gives us a measure of how well the actual outcomes are replicated by the model. This is based on the total variation of prediction explained by the model. R^2 is always between 0 and 1 or between 0 to 100 per cent.A value close to 1 for R^2 means a good fit.

To conclude the results, only ARIMA has the least value for MAE, MSE and RMSE. Other models cannot outperform Random walk. This result coincide with most of the literature. The prediction errors of TCNN are lower than those of TCNN with attention. This indicates that TCNN with attention cannot predict the value in the next period well.

6.1.2 DM test results for the selected models

Table 3.

Model	DM t-stat	P value
XGBoost	-3.818	0.000226
Elastic Net	-7.115	0.000060
SVM	-7.237	0.000122
Random Forest	-3.106	0.002424
Linear Regression	**-2.083**	**0.039515**
TCNN	-5.917	0.001340
TCNN with attention	-8.961	0.001340
UnSim LSTM	-4.429	0.015467
LSTM	-4.582	0.008433

Diebold-Mariano(DM) [19] test statistics for each model are reported in Table 3. We use it to determine whether forecasts made by the models are significantly different from random walk without drift. A statistically significant negative DM statistic indicates that the models do not have superior forecast accuracy rates relative to random walk without drift. The Linear Regression model has a least negative value which is better than other tested models.

6.1.3 Direction-of-Change Prediction Accuracy Rates for the selected models

Table 4.

Model	Prediction Accuracy Rates
XGBoost	45.37%
Elastic Net	51.85%
SVM	50.93%
Random Forest	48.15%
Linear Regression	**58.33%**
TCNN	43.52%
TCNN with attention	49.07%
UnSim LSTM	51.85%
LSTM	57.40%

The results in Table 4 shown the overall performance of each model. If the prediction accuracy rate is above 50%, that model may have a good prediction. On the other hand, if the prediction accuracy rate is below 50%, it implies that model has a bad prediction. So Linear Regression model has the highest prediction accuracy among all tested models while TCNN has the least prediction accuracy. Our proposed TCNN with attention performs better than TCNN. Elastic Net, SVM, UnSim LSTM & LSTM models have a result above 50 which indicates that they may have some forecasting capability.

6.1.4 Cumulative Profit comparisons for the selected models

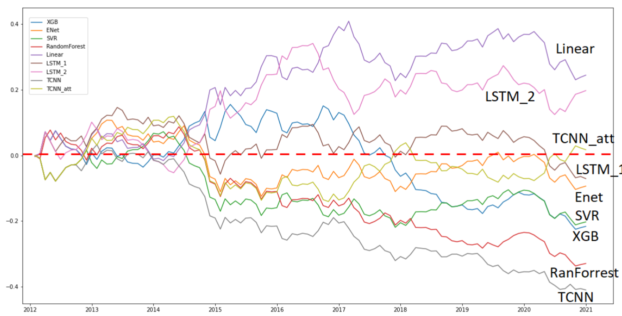

Figure 7.

We formed a virtual investment portfolio which long the EUR/USD at present period when a model predicts them to go up in the next period. Otherwise, short sell the exchange rate. From the results on Fig. 7 , it is observed that only TCNN with attention, LSTM and Linear Regression models have a positive returns of investment with the relevant predictions. All other models have a negative returns of investment. Surprisingly, Linear Regression generates the highest cumulative profits. LSTM ranks the second and this indicates that there may be a momentum effect since the previous exchange rates help to predict the next period value. Our proposed TCNN with attention ranks the third, which is far better than TCNN does. Thus, the attention layer helps to predict the direction of change and generate higher profits.

7. Conclusion and Future Work

With the vision of building a comprehensive and accurate US Dollar to Euro exchange rate prediction model, we not only collected 11 types of real financial data since 2003 from the market, but also applied various machine learning models from the elementary basic model like liner regression to the profound Temporal Convolutional Neural Network in this project. Furthermore, we evaluated the performance from different points of view including a set of common regression metrics, prediction accuracy, benchmarking with random walk model and cumulative profits generated from each model.

In Section 6, the prediction results indicated different performance among the models. More importantly, the results gave us an introspection about the necessary correlation between complexity of the model adopted and the performance. It is because the TCNN got lowest accuracy among the models while the simple model linear regression got the best accuracy rating. In this connection, the best cumulative profit is produced by liner regression model.

This project could be further extended to following areas. Firstly, complex models like LSTM and TCNN are powerful. The inferior performance of these models in this project are properly due to parameter setting. It is believed that further modification and improvement work on these models could produce more accurate prediction which outperform other basic models.

Secondly, in fact, currency exchange rate is not limited to US Dollar to Euro. Other currency exchange rates such as USD to JPY, USD to GBP, USD to CNY, USD to CAD, USD to AUD and USD to CHF also play important roles in the economic market. As such, prediction of the trend of aforesaid currency exchange rate is considered valuable.

References

[1] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo. Stock price prediction using the arima model. In *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, pages 106–112, 2014.

[2] S. Bai, J. Z. Kolter, and V. Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *CoRR*, abs/1803.01271, 2018.

[3] T. M. C. Amat and G. Stoltz. Fundamentals and exchange rate forecastability with simple machine learning methods. In *Journal of International Money and Finance*, pages 88, 1 – 24, 2018.

[4] C. L. C. L. Charles Engel, Dohyeon Lee and S. P. Y. Wu. The uncovered interest parity puzzle, exchange rate forecasting, and taylor rules. In *Journal of International Money and Finance*, pages Vol. 95, pp. 317–331, 2019.

[5] N. M. Charles Engel and K. D. West. Factor model forecasts of exchange rates. In *Econometric Reviews*, pages 1–24, 2014.

[6] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. *CoRR*, abs/1603.02754, 2016.

[7] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar, Oct. 2014. Association for Computational Linguistics.

[8] T. K. Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282 vol.1, 1995.

[9] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[10] L. Z. Kim, Y. J. and Tornell. Speculators' positions and exchange rate forecasts: Beating random. *Working Paper*, 2014.

[11] M. C. Marcos dal Bianco and G. Perez-Quiros. Shor-run forecasting of the euro-dollar exchange rate with economic fundamentals. *Working Paper*, 2012.

[12] R. Meese and K. Rogoff. Empirical exchange rate models of the seventies: Do they fit out of sample? In *Journal of International Economics*, pages 14: 3–24, 1983.

[13] V. Pacelli. Forecasting exchange rates: a comparative analysis. In *International Journal of Business and Social Science*, page Vol. 3 No. 10, 2012.

[14] D. Pradeepkumar and V. Ravi. Forex rate prediction using chaos and quantile regression random forest. In *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, pages 517–522, 2016.

[15] C. G. Rojas and M. Herman. Foreign exchange forecasting via machine learning. 2018.

[16] B. Rossi. Exchange rate predictability. In *Journal of Economic Literature*, pages Vol. 51, No. 4 (DECEMBER 2013), pp. 1063–1119, 2013.

[17] M. A. C. S. Ramakrishnan, S. Butt and H. Ahmad. Forecasting malaysian exchange rate using machine learning techniques based on commodities prices. In *International Conference on Research and Innovation in Information Systems (ICRIIS)*, pages p. 1–5, 2017.

[18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural In formation Processing Systems, (NIPS)*, 2017.

[19] D. F. X. and S. Mariano. Comparing predictive accuracy. In *Journal of business economic statistics*, pages 13(3), 253–264, 1995.

[20] A. G. P. Yin-Wong Cheung, Menzie Chinn and Y. Zhang. Exchange rate prediction redux: New models, new data, new currencies. In *Journal of International Money and Finance*, pages Vol. 95, pp. 332–362, 2019.

[21] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2003.

Appendix

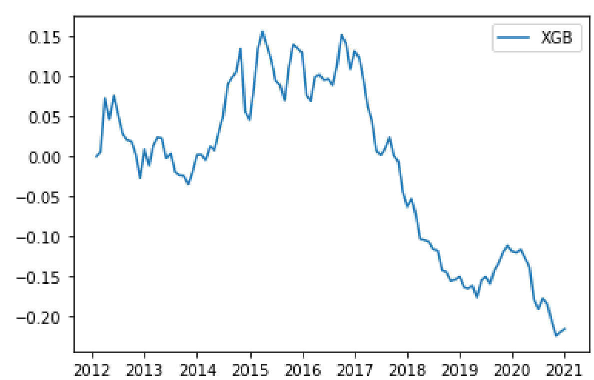

Figure 8. Cumulative Profit of Model XGBoost

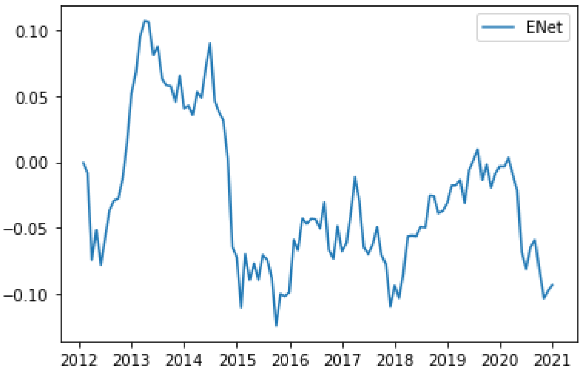

Figure 9. Cumulative Profit of Model Elastic Net

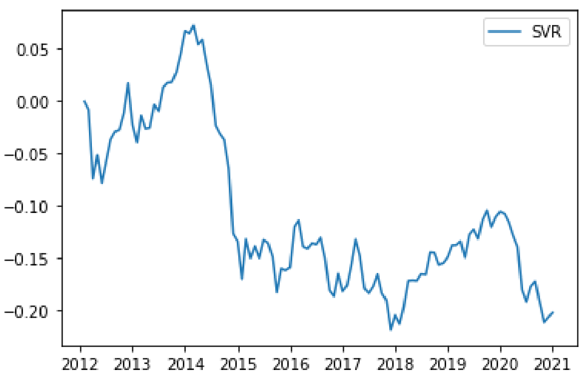

Figure 10. Cumulative Profit of Model SVM

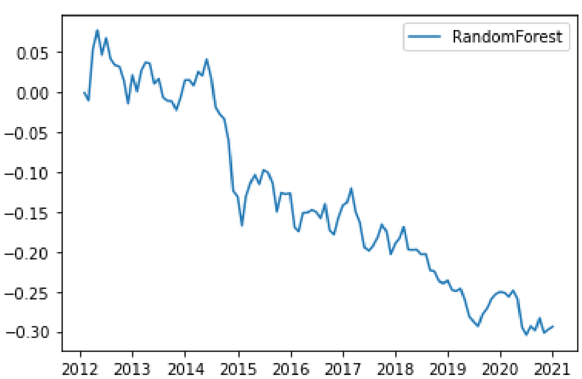

Figure 11. Cumulative Profit of Model Random Forrest

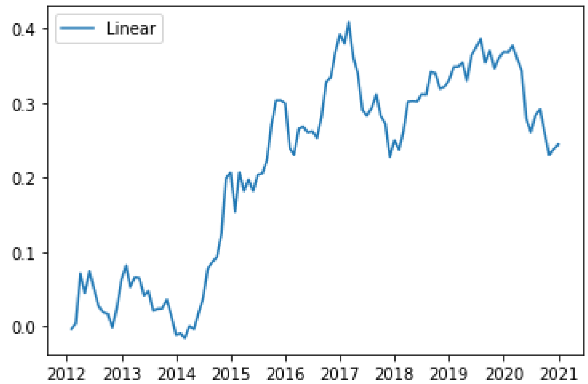

Figure 12. Cumulative Profit of Model Linear Regression

Figure 13. Cumulative Profit of Model TCNN

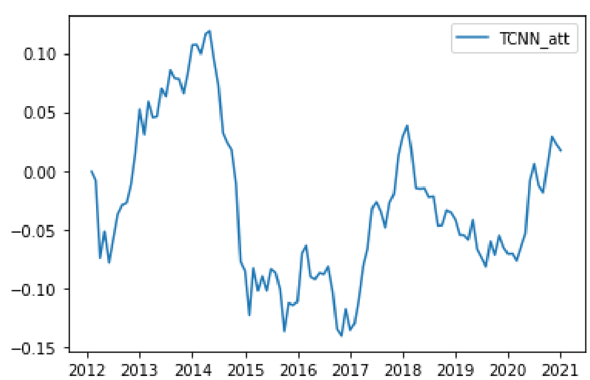

Figure 14. Cumulative Profit of Model TCNN with attention

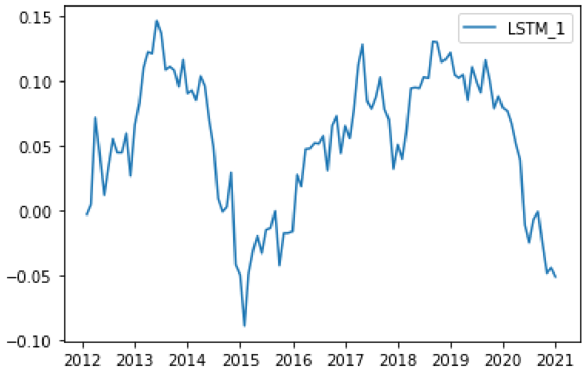

Figure 15. Cumulative Profit of Model Unsim LSTM

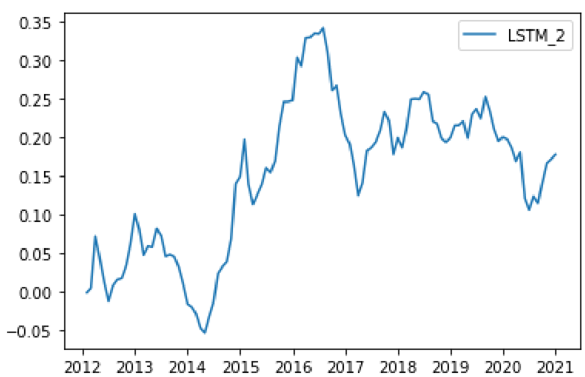

Figure 16. Cumulative Profit of Model LSTM