Currency Exchange Prediction

Yuntao Zheng
Carnegie Mellon University
Pittsburgh, PA 15213
yuntaozh@andrew.cmu.edu

Shaoxuan Zheng
Carnegie Mellon University
Pittsburgh, PA 15213
shaoxuaz@andrew.cmu.edu

Yujun Lu
Carnegie Mellon University
Pittsburgh, PA 15213
yujunlu@andrew.cmu.edu

Bin Peng
Carnegie Mellon University
Pittsburgh, PA 15213
binp@andrew.cmu.edu

1 Introduction

Predicting currency exchange rates is a crucial research problem with far-reaching economic implications. It plays a pivotal role in international transactions for businesses and investors, impacting financial stability and global economies. With accurate prediction, it will be convenient for predicting economic trends and investments which can provide help for future business decision. In this project, we are going to predict the exchange rate for various Western currencies, and see if it is possible to make a profit by exchanging our existing money to other currencies. More specifically, assumming we have 1000 USD, and we have access to currency converters for the following currencies only: USD ,EUR, GBP, CHF and we want to see if we can obtain profit at the end of 5 trading days. To make our project more accurate and flexible, we added multi-variables to predict the trend based on more features and facts.

2 Related Work

2.1 ARIMA

In 1970, George E. P. Box, e.t. published a classic book titled "Time Series Analysis: Forecasting and Control," [1]. The book provided a detailed introduction to AR and MA, which were already well developed. What makes it different is that the book first introduced ARIMA, which combined AR and MA to be a useful algorithm for time series analysis and prediction. By implementing ARIMA and combining it with prior knowledge in specific fields, people have successfully developed many predictors. For example, J. Contreras, e.t.[2] utilized ARIMA to predict next-day electricity prices. What's more, there are some similar jobs done by Babu, e.t.[3] based on ARIMA, which serves as the state of the art of ARIMA.

2.2 Prophet

In recent years, the application of advanced predictive modeling techniques, such as the Prophet model developed by Taylor and Letham[4], has gained prominence in the field of currency exchange rate prediction, proved by Fanoon and Nihla[5], which serves as baseline of Phophet. Prophet, an open-source forecasting tool initially designed for time series data, offers a robust framework for capturing the complexities of currency exchange rates.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

2.3 SARIMA

SARIMA model, as an updated and improved version of ARIMA, can adapt to more complicated datasets with seasonal components[6]. Based on large amounts of seasonal and time-sensitive data sets collected nowadays, it has been applied to different industries such as the research done by Amer Malki, e.t. regarding the prediction of the demands of the COVID-19 vaccine in the next few days[7]. SARIMA used 7 parameters where the first three parameters belong to the non-seasonal part and the last four parameters are designed to be seasonal, which will introduce the seasonal features of the data into the actual modeling and training process [7]. We also used Auto ARIMA to find the best parameters based on datasets. Auto ARIMA basically finds the optimized parameters for the SARIMAX model which provides a more accurate and scientific way to improve the performance of the SARIMAX.

2.4 Multivariate SARIMA

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model stands as an advanced iteration of the traditional ARIMA framework, uniquely suited to handle complex datasets that exhibit both non-seasonal and seasonal patterns. Its application has been widespread, proving particularly effective in areas where data is both time-sensitive and influenced by seasonal variations, such as public health and financial markets.

This paper delves into the use of a multivariate SARIMA model for forecasting the exchange rates of key currency pairs including USD/EUR, USD/GBP, and USD/CHF. The volatile nature of currency markets, influenced by various economic indicators and global events, presents a fertile ground for applying sophisticated statistical models.

In enhancing the predictive power of the SARIMA model for currency exchange rates, we incorporate a series of exogenous variables: gold price, oil price, and Dow Jones values. These elements are critical in the global economic framework, often exerting a substantial impact on currency valuations. The inclusion of these exogenous factors allows the model to not only grasp the inherent seasonal trends in exchange rates but also to reflect the wider economic dynamics that influence currency fluctuations.

3 Datasets

The basic dataset that our project use are from the yahoo website on the exchange rate for various currencies below:
1. USD/EUR: https://finance.yahoo.com/quote/USDEUR%3DX/history?p=USDEUR%253DX
2. USD/GBP: https://finance.yahoo.com/quote/USDGBP%3DX/history?p=USDGBP%253DX
3. USD/CHF: https://finance.yahoo.com/quote/USDCHF%3DX/history?p=USDCHF%253DX

The range of our data is from 2013-11-04 to 2023-11-03, which is the data for 10 years. The reason why we choose it to be 10 years is because currency exchange rate has seasonality, which is typically shown in unit of year. We choose to use close price and trading dates as our input data for each day. When downloading data from website, there are nan values for 3 days, we choose to manually drop them because they are not trading days. What is different than the handout is that, we only use the first 3 currency exchange rate data that involves USD, because other pairs can be infered.

Besides this, we have also opted other datasets of Macro-Economy:
1, Gold price: https://www.macrotrends.net/1333/historical-gold-prices-100-year-chart
2, Oil chart: https://www.macrotrends.net/1369/crude-oil-price-history-chart
3, Dow Jones Index: https://finance.yahoo.com/quote/

The timespan is the same as exchange data. Macro economy data plays a crucial role in exchange rate prediction as it reflects the economic health and stability of a country, which are key determinants of currency strength. For example, the fed fund rate is a key factor of how valuable USD is.

4 Experiment

4.1 Working Experiment

4.1.1 ARIMA

In our ARIMA experiment, we initially partitioned our dataset into training and validation sets. The validation set, consisting of only 5 samples, served the purpose of assessing ARIMA's capability to forecast next week's trend. Employing grid search, we determined the optimal hyperparameters for ARIMA before training the model on the entire dataset with these refined parameters. Ultimately, this approach yielded a commendable result, achieving a maximum profit of $0.9.

4.1.2 SARIMA

Similar to ARIMA, we partitioned our dataset into training and validation. And we used similar grid-search parameter tuning approach. The best parameter is (0,1,2), (0,1,2), this implies the first derivative of data is stationary. However, SARIMA does not work as expected and we will discuss later.

4.1.3 Multivariate SARIMA

In our extensive analysis, the second optimal model exhibiting the lowest Mean Squared Error (MSE) emerged as the multivariate Seasonal Autoregressive Integrated Moving Average (SARIMA). This sophisticated model surpasses the performance of our previous best-performing model, the ARIMA, identified during our midterm evaluation. The key enhancement in the multivariate SARIMA lies in its integration of additional influential variables, carefully selected based on their potential impact on currency exchange rates. These include the prices of gold and oil, alongside the Dow Jones Industrial Average value. Each of these variables plays a pivotal role in the global economic landscape, making them prime candidates for influencing fluctuations in currency markets. The incorporation of these variables allows the multivariate SARIMA to capture a more comprehensive picture of the market dynamics, leading to more accurate and reliable predictions. By embracing a more holistic approach, this model leverages the intricate interplay between various economic indicators, thus providing a nuanced understanding of the factors driving currency exchange rates. This advancement marks a significant step forward in our forecasting capabilities, offering a robust tool for navigating the complexities of the financial world.

4.2 Non-working Experiment

4.2.1 Prophet

Based on the prediction shown for prophet and the weight for the assembled model, propeht is not working as expected among the 3 models. Perhaps the reason is prophet is mainly for data with strong seasonality, but the currency exchange rate curve does not show strong stable seasonality trend. For future improvement steps, the output for prophet contains information includes yhat/trend_high/low.

5 Evaluation Metrics

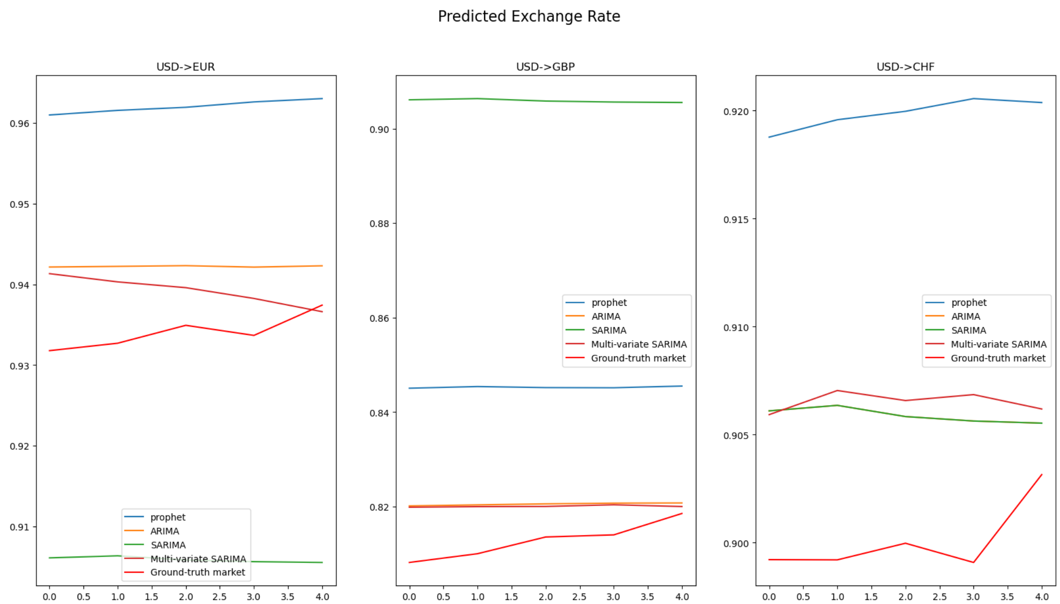
The metric will be Mean Squared Error (MSE) between the max profit computed via our model and the actual profit that can be obtained.

6 Result

According to our currency exchange predictions, it appears that we're set to gain $1.7 from Nov.$6^{th}$ to Nov.10^{th} by following the optimal policy: exchange USD to CHF on Monday, exchange USD to GBP on Tuesday and Wednesday. While this forecast is promising, it's essential to acknowledge the ground truth, which suggests a potential gain of $12 if we adhere to another policy: exchange USD to GBP on Monday and keep it. It's crucial to weigh both the predicted and actual outcomes to make informed decisions and ensure our strategies align with the reality of the market.

Table 1: MSE of Four Different Methods

	Prophet	ARIMA	SARIMA	Multi-variate SARIMA
MSE	108.3	142.4	157.5	123.7



7 Discussion and Analysis

7.0.1 Prophet

The Mean Squared Error (MSE) of the Prophet model is higher than average when compared to other models. This suboptimal performance could be attributed to Prophet's heavy reliance on seasonal trends in the training data. However, current exchange rate data may not exhibit significant seasonality, which could be adversely affecting the model's accuracy.

7.0.2 SARIMA

We assume that SARIMA could be better than ARIMA because of the seasonal nature of financial market. However, SARIMA turns out to be less optimal than ARIMA, even though we used the same grid-search parameter tuning approach. To explain this, we have some assumptions:

1, Exchange market data is no seasonal:
Exchange rate data is irregularly seasonal. If the data does not have a significant regular seasonal pattern, the additional complexity of the SARIMA model might lead to overfitting.
2, Complex factors lead both ARIMA and SARIMA underperform:
Exchange rates are influenced by a myriad of complex factors like geopolitical events, economic policies, interest rate changes, and market sentiment. Uni-variate models might perform indeed the same and the difference just comes from noise.

7.0.3 Multi-variate SARIMA

Multi-variate SARIMA is often regarded as the most effective model among the various options we have considered. Its strength lies in capturing complex seasonal patterns in time series data, making it a preferred choice for forecasting tasks. However, the predictive accuracy of this model might be limited in the context of currency fluctuation forecasting. Currency values are influenced by a multitude of factors, including geopolitical events and wars, which are challenging to quantify and incorporate into any statistical model.

These external factors, often referred to as exogenous variables, can have a profound and often unpredictable impact on currency markets. Unlike typical market data, events like political unrest or military conflicts do not have a regular pattern and are difficult to encode in a quantitative model like SARIMA. As a result, while SARIMA can capture and predict based on historical patterns in the data, its ability to foresee changes driven by these non-quantifiable factors is limited. This limitation underscores the inherent challenges in modeling and forecasting in the highly dynamic and complex domain of currency exchange rates, where external, non-quantifiable influences play a significant role.

8 Timeline

Tasks we have already accomplished:
Nov 1: Implementation of Three Models;
Nov 4: Combination of Models;
Nov 5: Submission of Midterm Report;
Dec 5: Implementation of multi-variate SARIMA
Dec 7: Draft of Final Report
Dec 10: Submission of Final Report

9 Division of Work

Yuantao Zheng: Data engineering,implementation of ARIMA and ensembling models, result visualization.
Yujun Lu: Data engineering, implementation of SARIMA and ensembling models.
Shaoxuan Zheng: Data engineering, implementation of Prophet, multi-variate SARIMA and ensembling models.
Bin Peng: Data preprocessing and implementation ARIMA.

References

[1] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control.* John Wiley & Sons, 2015.

[2] J. Contreras, R. Espinola, F.J. Nogales, and A.J. Conejo. Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18(3):1014–1020, 2003.

[3] AS Babu and SK Reddy. Exchange rate forecasting using arima. *Neural Network and Fuzzy Neuron, Journal of Stock & Forex Trading*, 4(3):01–05, 2015.

[4] Sean J. Taylor Benjamin Letham. Forecasting at scale. *The American Statistician*, 2018.

[5] Fanoon Raheem and Nihla Iqbal. Forecasting foreign exchange rate: Use of fbprophet. In *2021 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, volume 4, pages 44–48, 2021.

[6] Jason Brownlee. A gentle introduction to sarima for time series forecasting in python. *Time Series*, 2019.

[7] Amer Malki, El-Sayed Atlam, Aboul Ella Hassanien, Ashraf Ewis, Guesh Dagnew, and Ibrahim Gad. Sarima model-based forecasting required number of covid-19 vaccines globally and empirical analysis of peoples' view towards the vaccines. *Alexandria Engineering Journal*, 61(12):12091–12110, 2022.