

Article Title

MIGUEL SANCHEZ JEFF HSUING KYLE ZHANG

migchez@umich.edu test@umich.edu

Abstract

Forex, or the Foreign Exchange Market is a market for the trading of currencies from around the world. The system is highly complex and volatile, and thus, even though massive amounts of data on the market are freely available, prediction of the daily trends has been a difficult problem. By modeling Forex prediction as a simple binary classification problem (currency will increase/decrease in value relative to another), we are hoping to be able to leverage modern ML techniques and open sourced libraries in order to predict daily Forex trends with a high level of accuracy. We trained a Neural Network and implemented a Gaussian Process to predict different future foreign exchange values. Then based upon these different model values, we can deduce a predicted trend of currency movement.

I. INTRODUCTION

This section presents some background about the foreign exchange market and the reasons why our team is tackling this problem. First, the concepts of currency exchange rates and foreign exchange markets are explained. Followed by the motivation behind this project.

Currency exchange is the trading of one currency against another and usually takes place as follows[9]. Consumers typically come into contact with currency exchange when they have to convert their current currency in to the currency of a foreign country they intend to travel to. Businesses typically have to deal with currency exchange when they conduct trade outside their home country. Investors and speculators also trade currencies directly in order to benefit from the movements in the currency exchange rate. For example, if an American investor believes that the Swiss economy is strengthening, and as a result expects the Swiss Franc to appreciate in value relative to the U.S. Dollar, the investor may want to buy Swiss Francs and take what is referred to as a long position. Similarly, if the Swiss Franc is believed to go down, the investor may want to sell Swiss Francs against U.S. Dollars, thus taking a short position. Those transactions are often conducted to take advantage of movements in the market over very short time periods. Commercial and investment banks also participate in the currency market, as well as governments and central banks when they try to intervene for adjusting economic imbalances. Currency exchange rates are always quoted for a currency pair using ISO code abbreviations[1]. For example, US/JPY refers to the two currencies US Dollar and Japanese Yen. The first is the base currency and the second the quote currency. Therefore, US/JPY specifies how many US Dollars you get for one Japanese Yen.

The currency exchange market, also referred to as the Forex market, was established in 1971, when floating exchange rates began to materialize[9]. In terms of trading volume, Forex is the world's largest market, with daily trading volumes in excess of \$4 trillion USD[1]. It is by far the most efficient and easily accessible market. Today, an overwhelming majority of all Forex transactions involve a few major currencies, the US Dollar (USD), Euro (EUR), Japanese Yen (JPY), Swiss Franc (CHF), and British Pound (GBP). While there are other currencies in the world, we will primarily be only interested in these 5. The Forex market is open for business around the clock 24/7.

The reason why we want to predict the exchange rates in the Forex is due to the availability of data and the potential financial and social gains. As we have stated before, there are massive amounts of data recorded for this problem. This is due to the accounting principles of the world in order to audit trades of currencies in order to safeguard against fraudulent practices in the market. Due to these accounting principles, there is data available at such a minute scale that it becomes ineffective. In the previous paragraph, we discussed that there are several factors that can influence the exchange rates. Most countries have detailed historical data on several economic benchmarks which we can use for our problem.

The incredible range of data available to us makes this problem attractive to solve. As for the potential financial and social gains, if we can predict the increase/decrease of a currency, there is obvious financial gains in anticipating exchange rate changes. For example, investing in a currency that will rise in value relative to another, and then exchanging the currency after the rise has occurred will allow an investor to increase their assets. The less obvious social benefit to creating a more accurate foreign exchange rate predictor lies in the fundamentals of the market. Due to the effects of information asymmetry, where people don't have the same information when making trades, some people can take advantage of other's ignorance and benefit at the cost to others. In practice, investors with more resources, such as the institutional firms, are more informed than the average investor and thus can take advantage of that fact. So if we can create a predictive model that works, we can level the information playing field for the average investor and and save them the costs of information asymmetry.

However, if we do somehow engineer a predictor of exchange rates with high enough accuracy on historical data, there is no guarantee that it will work on future/current data. This is because of the anticipatory behavior of participants in the market. This behavior is when participants execute actions that benefit themselves based upon information or beliefs on what the market will do. So if we engineered a perfect predictor, it will immediately become redundant as all participants in the market will trade on this data and immediately remove all opportunities the information can create or exacerbate the volatility in the market.

II. PROPOSED METHOD

Our proposed method pipeline is engineering our data, find the optimal feature set using a small data set Gaussian Process, cluster input data to train separate Neural Networks, then predict a exchange rate value with full Gaussian Process and trained Neural Network some x days into the future. Finally based on predictions, we classify if the exchange rate will appreciate or depreciate. This section will detail in order how we go about implementing these steps. All implementation code was programmed in python.

Firstly, there has been much research into how various factors influence the exchange rate. There is also much debate over which factors are the causation of changes in exchange rate and which factors are merely correlated. From our initial research into related work, we have found that there are some widely used and accepted factors and these are the ones that we will use as features [3,4,5, 8]. They are the inflation rates of countries, money supply of the currency, current accounts (trade balance), GDP, interest rate, and stock indices. We also need the historical spot, closing, opening, and volume data for the US/JPY rate on the Forex market for input and output data. After gaining access to several databases such as Bloomberg and Factset, we have acquired the relevant information but each of these features and outputs are a time series on some frequency that varies. So, after joining the data to a common frequency, we can now use these as our data for our machine learning models. Another transformation we are doing to the data is to take the log of exponential growth metrics such as GDP, current accounts, money supply to give us a linear pattern to data. Unfortunately, upon further investigation, money supply, interest rate, and current accounts are calculated and recorded monthly and sometimes quarterly. The granularity of this data is too low to be used with our other data and thus had to be discarded. Another problem we encountered was the volume feature. Volume is the amount of trades (currency swaps) that occurred on that day and gives no indication of movement of currency. So we calculated several momentum metrics using volume and prices like RSI[12], moving averages[11], OBV[13] which can tell us how fast or long price changes might occur or last. So, after engineering our features with these two transformations, eliminating infrequent ones, and constructing more meaningful ones, we have a usable dataset of 16 features for the problem.

However, even if these factors are widely used, the fact that the problem has remained unsolved may mean that these factors may not actually capture the complexity of the exchange rates or are not relevant to the problem. Rather than accepting these factors blindly, we will do some feature testing to find the optimal set of features. Our proposed method was to brute-force every combination of features using our implementation of Gaussian Process on a small dataset to quickly get results. Our Gaussian Process uses a

squared exponential kernel for the covariance matrix. In this initial GP we predict the next day exchange rate t_{N+1} by calculating the conditional distribution $p(t_{N+1}|t_N)$. After predicting values using all combinations of features, we choose the set that minimizes the RMSE.

After finding the optimal features, we are proposing to apply clustering techniques like K-means and Gaussian Mixture Models to help segment the data. Then we can train a separate neural network for each cluster of data. The motivation behind this is that we can have more specialized neural networks for specific sets of data. This is superior to training one neural network on the entire data because trying to fit all data in one network could result in underfitting problems. With several neural networks trained on clustered data, we can have more fitted neural networks to predict all data without worrying about the problem of overfitting. So when it comes to testing, we can cluster the test input into its appropriate cluster then predict using the trained neural network of that cluster. We believe this approach will allow us to better understand the volatility in the market.

After clustering, we make a prediction of x days in the future about the exchange rate. 5. As the day-to-day exchange rate does not fluctuate dramatically, we attempted to predict beyond just one day into the future both as an academic exercise and to seek potential for greater capital gains (from longer changes in the exchange rate). In the Gaussian Process model, this is done by calculating the conditional distribution for not only the target value (The USD-JPY exchange rate), but also all the features used. As specified in the feature selection paragraph, our Gaussian Process uses a squared exponential kernel for the covariance matrix. The GP predicts the next day exchange rate t_{N+1} by calculating the conditional distribution $p(t_{N+1}|t_N)$. By switching out the target values t_N for the training set with values of another feature f , we can create a new conditional distribution $p(f_{N+1}|f_N)$ and make predictions for all the features involved at each data point. By predicting all the feature values f_{N+1} , we have synthetically generated a new data point which maps the set of predicted feature values to the predicted exchange rate. From this we can expand our existing covariance matrix such that the new expanded covariance matrix C_{N+1} includes the test point X_{N+1} . Now to predict 2 days ahead, the covariance matrix will be further expanded to C_{N+2} and the Gaussian Process is expanded to model the joint distribution of the training data along with the test point X_{N+1} . This process can be repeated to predict into the future as we synthetically generate test points $X_{N+1}, X_{N+2}, \dots, X_{N+k}$. The Gaussian Process was optimized with hyperparameter values identified using maximum likelihood as given by the equation in Bishop equation 6.69 [10 pg 311]. The parameters were the the sigma value (width) of the squared exponential kernel and beta, which was the precision of the noise in our data.

As for the NN, we have a basic multilayer forward feedback perceptron neural network implemented. The activation function of each neuron is a tanh function. Since we are working with time-series data, we must structure our inputs into patterns in order to impart some notion of time in the neural network. We must always structure our features so that there is order to the inputs. For example, if we had samples $x_1, x_2, x_3, x_4 \dots x_N$ and their outputs $o_1, o_2, o_3, o_4 \dots o_N$ respectively, we train the neural network by using inputs x_1, x_2, x_3, x_4 to predict o_5 . Then to predict o_6 , we use x_2, x_3, x_4, x_5 . With respect to the details of our project, we use the past 5 days, in order, as input to predict today's exchange rate. By keeping this pattern of inputs, we can construct inputs that can reasonably model the flow of time. We are also tweaking the hyperparameters such as number of hidden layers or neurons at each layers to experiment with their effects on accuracy. By using a script to test different combinations of hyperparameters, we can compare errors and use the optimal ones. The ranges for hidden layer size varied from 1 to 5, one step at a time. Hidden units in each layer varied from number of features to 100 with steps of ten at a time. The tool we are using to create these neural networks is neurolab in python.

Finally, after a prediction is generated for a day from the NN and GP, we compare it with the previous prediction to see if there is an increase or decrease and output +1 (increase) or -1 (decrease) for that day. We compare these to our test data to see how high our accuracies are. Our whole approach is that we are trying to simplify the problem down to a classification problem (increase or decrease from previous time periods). By doing this, regardless of our accuracy at predicting exact values of exchange rate, if we can predict which way the exchange rate will go for a future date, we can still achieve our goal. Unfortunately, during times

when the actual exchange rate does not change by much, our proposed solution will be very susceptible to error when classifying. Even if our error rate is cents on the dollar, if the exchange rate only moves by a cent, we can predict a wrong exchange movement even though we are fairly accurate to the actual exchange rate.

As for our approach for predicting x days in advance for GP, aside from the academic challenge, exchange rates are easier to predict as we extend the time horizon. Short term exchange rates are incredibly difficult to predict and no model consistently outperforms a random walk [14]. If this is true, then we should see better predictions as we increase the number of days ahead we predict.

III. RELATED WORK

As we stated in the previous section, there have been many studies that examine this specific problem. Listed in our reference are 3 of the most extensive, advanced, and understandable studies we have encountered in our research. Existing methods range from simple mathematical models like linear or polynomial regression[8] to extensive neural networks[4] and even methods we have never heard of before (Generalized Autoregressive Conditional Heteroscedasticity for example [3]). One similarity all of these studies share is that they try to predict the exact value of the currency in some future time interval. Another similarity the studies share is that the accuracies are not impressive. When the predictive models work, they always seems to have a time offset which prevent them from being useful. By time offset, we mean that the actual exchange rate seems to move in the same direction but sooner in time than what the predictive models indicate [4]. This can be easily explained by the inefficiencies in the marketplace like anticipatory behavior explained in the introduction, insider information, or quantitative policies of the government. For example, if some people in the market had insider information, or information not available to the public, that informed them where the exchange market would move, then they can take actions before everyone else is informed in order to benefit when everyone else finds out. These are all dynamics in the market that we must consider when we interpret other's and our own results.

Due to the fact that quantitative models of Forex prediction do not perform satisfactorily, some researchers and firms are investigating psychological factors that affect the exchange rates. Many companies like Thomson Reuters[6] and Bloomberg[7] have created analytical tools based on machine learning algorithms that determine what the general public's opinion (positive, negative, or mixed) is of a topic/stock. While we lack the knowledge, time, and resources to implement such a sentiment analysis model, we believe that factoring in the sentiment can help us improve the accuracy of our predictive models in the short run, specifically the volatile spikes in value. We believe that the temporary volatile spikes in a currency are caused by emotional and psychological reasons in participants in the market and this artificially inflates or deflates a currency. In the long run, the value of the currency is unaffected by these short term factors and is more dependent on the stability and security of the country that backs the currency[9, 14]. As stated in the Proposed Method section, we plan to cluster our data in order to train a Neural Network on each individual cluster. From the research we have done, we have not seen this technique employed. Combined with the data clustering technique, we believe this approach will help increase the accuracy of existing and new models.

IV. EXPERIMENTAL RESULTS

From our results, we were fairly successful at all in predicting true values of exchange rates but unsuccessful at predicting if the exchange rate would appreciate or depreciate.

The clustering part of the approach was a failure. When the clustering algorithms ran, the generated clusters contained continuous blocks of data that were equal fractions of the entire dataset. In hindsight this result should have been obvious. Because we are working with a time series, the most important factor of our feature set would be the timescale so clustering algorithms will cluster within a time interval. From an intuitive standpoint, this would make sense since current exchange rates are usually close to its past rates and

any large scale changes would take time to occur. Because we were only sorting data points into equally-sized temporal clusters, there was no significance in doing so and we eliminated this extra step in our pipeline.

From our brute force feature selection script, we eliminated some features to go from 16 to 11 relevant ones. Most of them are exactly what we expected (GDP, closing, moving averages, momentum indicators). Out of curiosity, we tried this on features we did not expect to be relevant to the USD/JPY exchange rate, such as countries stock indices that were not US or Japan. Our results came back and showed that the Canadian Stock Index was part of the optimal feature set. However, logically the Canadian Stock Index should have little impact on the US/JPY exchange rate but it does. This shows us that there might be other well correlated features that we are using but are ultimately noise/irrelevant to our problem.

For GP’s performance, in Appendix 1, we see that our prediction values are relatively close in the first few days and then sharply veers off in accuracy. While the confidence interval appears to be very thin in the various prediction points, we clearly failed to accurately predict the exchange rates. This indicates that while we have training points that are very close/similar to the test point in question, they are actually mapped to very different target values (exchange rate). This suggests that we were not using the best set of features which truly captures the factors that influences the exchange rate.

Another phenomenon we can witness is that our prediction values slowly get close to the actual as we try to predict further into the future. However we only tested up to 4 days ahead, and the movements in the graph are absolutely not conclusive without further testing. Unfortunately, the complexity of the simulation increases with the number of days in the future we predict and our time constraint only allowed us to simulate 4 days ahead.

TODO generate NN graphs of prediction values and RMSE The best hyperparameters to use after running the combination script was a hidden layer size of 3 with the number of neurons being the number of features on the left-most and right-most layer. The number of neurons for the middle hidden layer was 50. For NN’s performance, in Appendix 2, we that our prediction values are much closer and follows the actual much closer than GP.

As for the prediction of next day movement, we can see in Appendix 3 that these were not good results. Because we are doing a binary classification, 50% correct would be the performance of a completely random model. We are getting TODO get classification accuracy of NN and GP and draw some conclusions from it

V. CONCLUSION

Even with fairly accurate prediction values, we are unable to classify consistently correct trends in the USD/JPY rate. Due to the price change sensitivity, even a very small error in prediction value can result in wrongful classification. Throughout the course of this project, we have learned how to use different Machine Learning libraries to actually implement a large scale problem. From feature set selection to hyperparameter testing, we implemented several scripts that can help with tweaking machine learning models. From our results, we show the difficulty in predicting exchange rate movements, with us not being able to consistently beat a random walk in accuracy. However, these results reinforce the idea in financial research that short term interest rates are incredibly difficult to solve [8, 9, 14].

REFERENCES

- [Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.