TIME SERIES FORECASTING: ANALYSIS OF LSTM NEURAL NETWORKS TO PREDICT EXCHANGE RATES OF CURRENCIES.

Concordia University

FRANÇOIS DAVID - 40046319

June 21th, 2019

SUPERVISOR: DR. THOMAS FEVENS

Final report presented as part of the requirements for the completion of the course:
COMP490 - Computer Science Project I - Summer 2019

1 ABSTRACT

The modern financial and economical market is now formed of multiple efficient and complex systems. A few techniques and theories were introduced in the last few decades that have revolutionized the understanding of those systems in order to predict financial markets based on time-series analysis. However, none have been proven to efficiently work on a consistent basis. In this project, I forecast the foreign exchange rate of currencies using a special type of Neural Network Modelling; LSTM. This type of modelling has been popular in many different applications for forecasting since it can identify complex non-linear relationship between variables and the outcome it wishes to predict. In comparison with the stock market, the exchange rates seem more relevant because of available macroeconomic data that can be included in the network so that it learns the effects of specific indicators on the rate to be predicted. The data is obtained via Quandl, which is an economical and financial platform that provides quantitative indicators for many countries. The model is then compared with an Exponential Moving Average and an Autoregressive Integrated Moving Average on three different metrics. Then, I compare and validate the model's ability to predict accurately future values and debate which of the models predicted the most accurately. To access a rich set of libraries, the programming is done in Python and all sample codes can be found on GitHub, accessible through a link printed on the last page of this report.

CONTENTS

2 INTRODUCTION

The economic and financial ecosystems are extremely complex and difficult to analyse. A specific stock on the modern market can be influenced by a lot of particular factors and situations that are not always made public information. Since some of that information is not available for this study, I will focus on forecasting the foreign exchange currencies which have relatively more available data from government reports and macroeconomic indicators. To predict future values, it will use modelling tools of machine learning to learn the trends and the appropriates correlations between indicators and the rate to be predicted that could be useful for the prediction. In general, the exchange rate can be seen as the value of another currency that is considered equal to one unit of the base currency. The bank of Canada and the government does not actually intervene to keep the exchange rate fixed, however it can influence it by printing money or changing the interest rate to keep it stable in the long-run but does not have the right to directly regulate it [1]. The foreign exchange rates are influenced by a multitude of economic factors, which we can integrate those into our model since the online database Quandl provides quantitative index representations for a large quantity of countries. They will help the model figure out tendencies between the variations of the foreign exchange rates and those economic indicators, which will help the model get a better understanding of the market and ultimately make better prediction. A big challenge is to format this data accurately so that the noise of one indicator does not outweighs the others only because the indicator is larger in magnitude. This problem and the factors chosen will be addressed in section 4. Ultimately, the model is compared with more common financial and time-series forecasting techniques. The implementation of the model is done with TensorFlow, a neural network open-source library designed by Google for the experimentation of deep learning with python.

3 BACKGROUND INFORMATION

3.1 RNN - Recurrent Neural Networks

I knew that a neural network would be adequate for this type of forecasting but there is a significant amount of different architectures that can be used, some more appropriate than others. My specific problem is to forecast a time-series therefore, the neural network needs to have the ability to predict based on the previous information that was passed to the model. An appropriate type of network would be the RNN (Recurrent Neural Network) are really powerful algorithms because they have internal memory cells that allows accurate predictions of sequential data [2]. Those memory cells allow the network to remember crucial details about the past sequences they received which allows them to predict accurately the next sequences of data, something that other algorithms cannot do [2]. This type of architecture has this special ability because of a feedback loop that passes information to adjacent cells on the same layer so that it keeps important features that are useful to the computation in the short-run. RNN are especially useful to predict short term dependencies that will give a direction to the prediction [2].

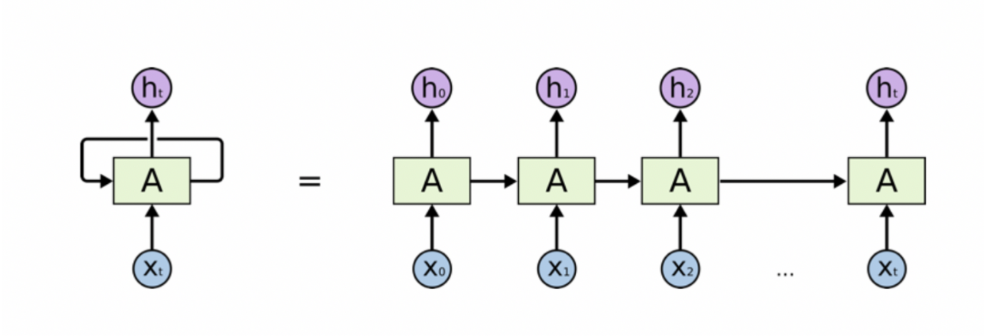


Figure 1: Recurrent Neural Network Cell Architecture [3]

The diagram above shows a typical structure of an RNN. In this example, Xt is the inputs to the network and form the input vectors to that specific layer. The output is denoted by h(t) and it is not necessarily the end of the network as it can have multiple hidden layers. The sideways output of each of the cells denoted by A only feed information to other cells on the same layer so that it keeps information about the previous steps used on the network. This type of network as a lot of promises in theory, but the applied version of RNN encounters some problems. In fact, the training of RNN uses feedback loop and long-term dependencies results in very large updates to neural network model weights which are sometimes non-computable [4]. The data transmitted to the other neurons is multiplied with the current input which leads to extremely large or small values, which results in the loss of the information that was meant to be preserved. It is because of the accumulation of the error gradients during the updates which will make the network unstable [3]. Through the training of the model, the updating of the weight becomes problematic because the gradient becomes so extremely large that it leads to an overflow if the value of the gradient is larger than 1 and it becomes 0 if it is smaller than 1 which leads to the network loosing the needed information to diagnose dependencies of previous calculations. For the purposes of this project, a specific general RNN would be exposed to that problematic and I judge that it might not be the most appropriate architecture.

3.2 LSTM – Long Short-Term Memory

In order to solve the drawback of RNN with long-term dependencies, I decided to use a LSTM, which is a specific type of recurrent neural network that attracted a lot of attention in the past few years in the artificial intelligence community. They have been introduced by two computer scientists in 1997, Hochreiter and Schmidhuber, and they were specifically designed to solve the vanishing or exploding long-term dependencies of recurrent neural networks [4].

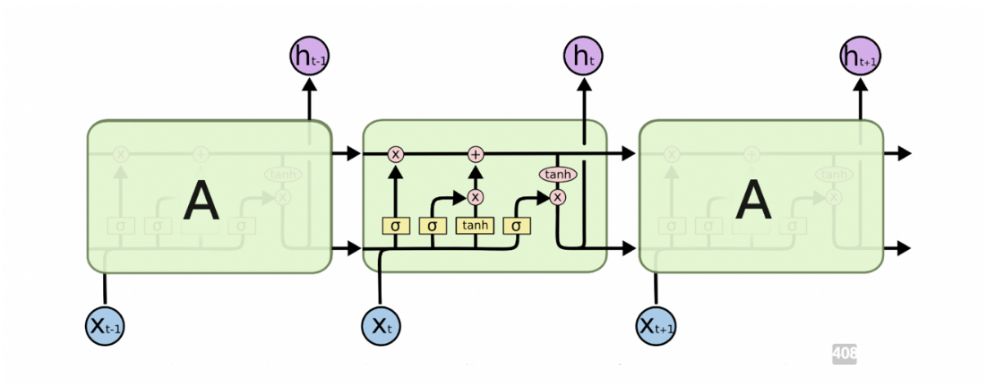


Figure 2: LSTM Cell Architecture [3]

As shown in the above diagram, the LSTM neuron have the ability to add, maintain or remove the information that is passed through the feedback loop. This remarkable ability is due to the gate made of a sigmoid neural net layer that optionally let through information depending on the value of the input [4]. This sigmoid layer outputs a value between 0 and 1 to tell how much information it should let through, which therefore regulates how much the information kept is preserved depending on the new input and the values from the previous iterations [4].

4 MACROECONOMIC INDICATORS INFLUENCING EXCHANGE RATES

Many economic indicators can be correlated with the value of the foreign exchange rate. In this study, I used a few metrics which are usually helpful in forecasting the exchange rate. The goal is that the LSTM will update its weights trough back-propagation to learn indications of future changes in the foreign exchange prices. For example, if the model learn that a specific indicator increases a few trading days before a decrease in the foreign exchange rate concerned, then it will adjust its model to learn from this tendency and increase the accuracy of the model in the future. None of the following indicators can predict alone the foreign exchange rate but it is a combination of a lot of macroeconomic indicators [5]. I used the GDP, the imports and exports, the inflation rate, the consumer price index, the interest rate, the government spending and the unemployment rate of the respective countries.

GROSS DOMESTIC PRODUCT The gross domestic product is the dollar value of everything that was produced in one year. Changes in the GDP reveals changes in the economic growth, which can directly impact the value of a country's currency [6]. In fact, an increase in the GDP demonstrates that there is an increase in the production. Therefore, it would translate to an increase in the demand for the appropriate currency. However, this increase can be counterfeited by inflation, which will be covered later in this section.

INTEREST RATES The interest rate of a specific country influences greatly the foreign exchange rate since the value of the National Interest rate will determine if the amount of foreign investment. Higher interest rate will normally increase the demand for the currency and increase the exchange rate [5].

INFLATION RATE AND CONSUMER PRICE INDEX The inflation rate is an important factor since it represents the amount of money printed by the country. Therefore, if it is increasing, the government is usually printing more money, increasing the supply of the currency.

IMPORTS AND EXPORTS Naturally, the imports and exports of each country will have a substantial effect on the exchange rate of the two currencies. When the imports of a country rises, there is more supply of the according currency (decreasing its price). When the exports rises, there is more demand for the currency and the price will increase.

GOVERNMENT SPENDING The government spending increase the income of the population, and should increase the demand for imports. This higher demand for goods outside of the country should lead to a depreciation of the currency [7].

UNEMPLOYMENT RATE The unemployment rate results will alter the exchange rate because a decrease in this indicator will increase the amount of income in the country and thus increase the demand for goods outside of the country. Therefore, it should lead to a depreciation of the currency [7].

5 METHODOLOGY

5.1 Data Sets

The data is acquired from the API Quandl. It gives indicators from a large number of economic indicators for almost all countries in the world. In order to evaluate the model on a different number of datasets, I decided to predict 3 foreign currency exchange rates, the USD/CAD, the GRB/USD and the AUS/USD. However, the data have to be modified in order to be adequate to fit in the LSTM model. Otherwise, if one indicator is considerably larger, more importance will be given to the weights accorded to their position since a small change could have much more effect than other indicators that might be as much important. For much of the data acquisition, I made a function to return only the percentage change from one entry to the next. Therefore, changes in a corresponding indicators' weight will change proportionally to its size. Also, since the frequency of some of the indicators is not daily, as the exchanges rate is, I filled the entries that had no data with zeros, since no changes as been recorded. The data covers the time period of approximately 26 years, from January 1st, 1993 to the March 29th, 2019. I separated the data so that the first 90% of the data will be used for the training of the model and the remaining 10% will be used for the validation/testing of the LSTM. I decided to use only 10% of the data to validate it since there are a large number of observations (approximately 6000) that can be taken for each data set and it will be easier to analyse the graph of the predictions of a smaller validation set (around 300 trading days). The foreign currency exchange rates acquired correspond to the following table.

Exchange Rates									
	USD/CAD			GRB/USD			AUS/USD		
	All	Train	Test	All	Train	Test	All	Train	Test
Amount	6790	6111	679	5565	5008	556	6412	5770	641
Mean	1.261	1.256	1.309	0.608	0.607	.618	0.768	0.743	0.842
Median	1.29	1.27	1.308	0.620	0.622	0.613	0.757	0.756	0.760
Max	1.61	1.61	1.457	0.731	0.731	0.812	1.103	1.103	0.811
Min	0.92	0.92	1.210	0.474	0.474	0.583	0.484	0.484	0.686

Table 1: Basic statistical indicators on the 3 different data sets.

There is a difference between the data set sizes is because some of the data was not available exactly from the exact same starting date but they are pretty similar. For the most part, the data seems adequate. However, we can see that the training contains more drastic changes since the all of the have the Max and Min of the overall data set inside of the training data.

5.2 Data Preparation

I had to prepare the data in order that all of the fluctuations in the indicators might be taken appropriately comparably at the beginning of the training of the model. For instance, if there is a change in the GDP from a few billions, at first I don't want the model to give much more importance to this indicator just because the value on the indicator is much higher. In order to solve that problem I normalized the data according to its range on the values between 0 and 1 according to its range of values. Moreover, I realized that the indicators that had inflation in them should be manipulated more since it will boost the values of the indicators without really indicating more economical strength. In order to make those indicators adjusted to the inflation, I changed the data so that it represents the percentage change from one day to the next and I then normalized the data set on the range of percentage changed in the period covered on the range of values between 0 and 1 like the other indicators that are already percentage values.

5.3 Evaluation Metrics

The results are evaluated according to three metrics: the RMSE, the MAE and MAPE. The RMSE stand for the square root of the mean of the squared errors. The MAE corresponds to the mean of the absolute value of the errors. The MAPE stands for the absolute value of the average percentage of the errors.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \hat{X}_i)^2}{n}},$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \hat{X}_i\right|$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right| * 100$$

All three of those metrics are commonly used in the analysis of time series prediction and indicate complimentary factors. For example, the RMSE is a lot more sensitive to large outlier prediction because it uses the square of the error. On the contrary, the MAE gives all the differences between the prediction and the target to be weighted the equally. The MAPE is useful since it re-scales the difference based on the percentage of the actual value to be predicted and can therefore compare the results on different data sets.

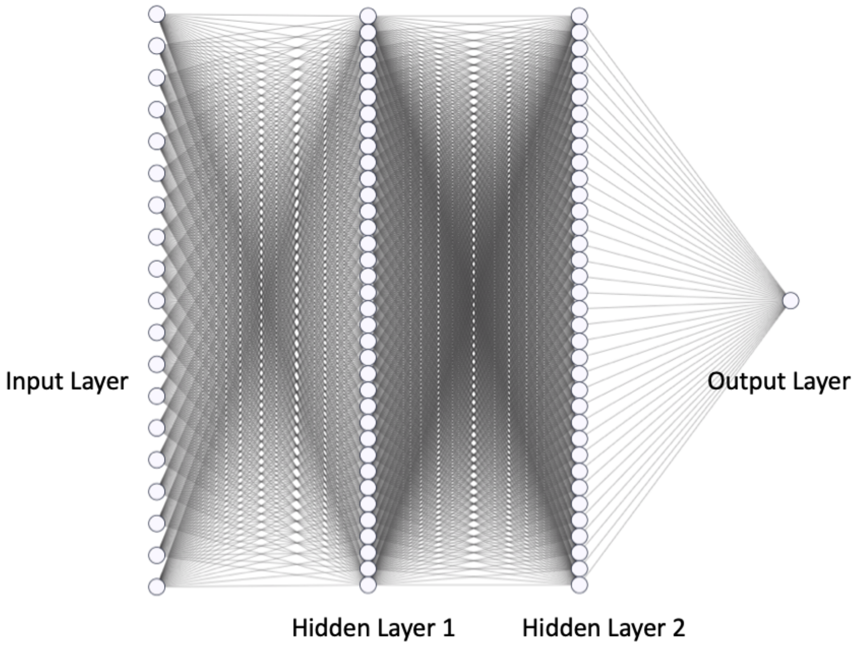
5.4 Hyperparameters' Optimization

I first needed to know how many hidden layers would my model use and how many cells would each of them contains. Generally, a network with 2 hidden layers have been found to detect complex features [8]. I also found a formula that is considered a rule of thumb when deciding for the amount of cells to include in each hidden layers.

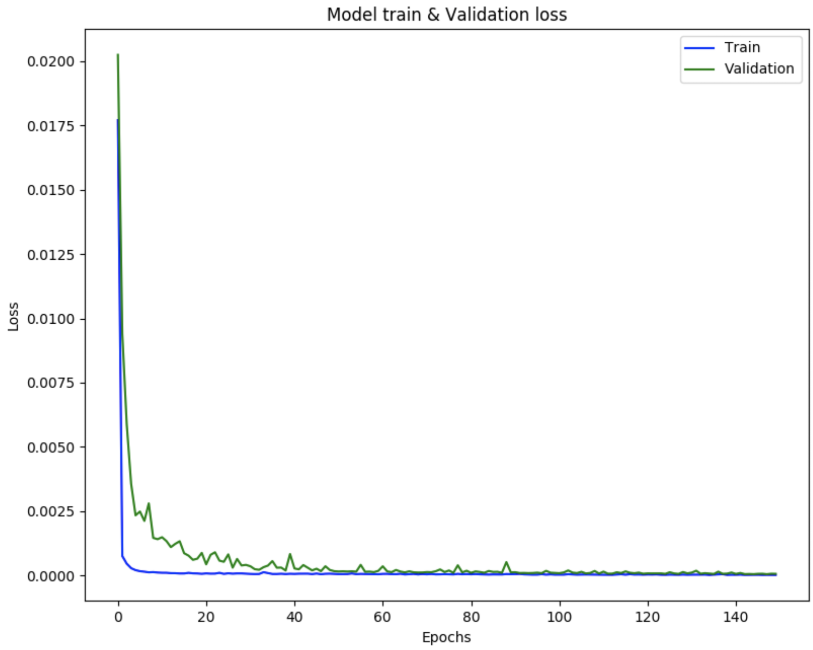
$$N_h = \frac{N_s}{(\alpha * (N_i + N_o))}$$

Figure 3: Formula for the amount of cells to include in the hidden layer [8].

In the formula above, N_h represents the number of cells to be included in the hidden layer, the N_S is the number of samples that will be passed to the network as training data, N_i is the number of inputs neurons, N_O is the number of output neurons and the value **a** correspond to a scaling factor that can range from 2 to 10. In the context of this project, the number of samples (N_s) in the training data is roughly 6000 sample days. The number of input neurons (N_i) to the networks is roughly 19 values (9 indicators for each country and the exchange rate value). The output neurons of the neuron (N_O) is only one value, the value of the projected rate. I tried the different values of **a** in the range of 2 and 10 on the 3 datasets and got the lowest RMSE for the value of 8, which leads to a number of cells of 37.5, which I approximated to 38 cells per hidden layer. Therefore, the chosen architecture looks like:



In order to determine if the LSTM model is not underfitted or overfitted. Underfitted being that the loss over the training data and the validation data are still decreasing and the model could learn from more epochs (iterations of learning over the training data sets that adjusts the weights of the algorithm) [9]. Overfitting is when the loss of the training data is still decreasing while the loss over the validation data is increasing [9]. The model performs well on the training data but not on the validation data meaning that the algorithm adapted too much to the specific cases of the training data which decreased its understanding of the overall trends to predict. One way to make sure that the epochs is appropriate for the learning is to analyses variations in the training loss and the validation loss. In order to be a good fit, where the model would perform well on both the training data and the validation data, both losses have to decrease and stabilize at approximately the same point [9]. Through trial and error, I found that 150 epochs were an appropriate number of epochs since both losses appear to stabilize at around the same point for all three data sets. For example, this is the graphing of the losses of the USD/CAD data set.



We can see that the training loss starts higher than the training loss, as expected, and goes down a bit more slowly than the training loss as the epochs increases. At around 30-35 epochs, the validation loss is similar to the training loss but some considerable oscillations occurs from 40 to 100. From the epochs 100 to 140, we see that those oscillations goes down and the curves are almost identical from 140 to 150. Since they seem to stabilize at that point, I decided to stop the training and concluded that the model was appropriately trained.

5.5 Exponential Moving Average

One of the comparison schemes chosen is predicting the next day with a exponential moving average. It consists of predicting the value of the exchange rate of the day to predict by calculating the arithmetic weighted mean of a certain number of previous values, the rolling window-size. The fact that makes the Exponential Moving Average more accurate to the Simple Moving Average and the Simple Weighted Moving Average is that it gives each value of the window of calculation an exponential weight so that the most recent weights are much more important [10]. It is more appropriate for financial forecasting since if a price change direction quickly, the values of the earliest data in the window will not be really representative of the current situation and the EMA exponentially diminishes the impact of old irrelevant values to prioritize recent ones [11]. The formula used for the calculation of the Exponential Moving Average is:

$$F_{t+1} = (X_t - F_t) * (2/(\text{period} + 1)) + F_t \tag{1}$$

In this formula, F_{t+1} represents the predicted foreign exchange value of the next trading day and X_t represents the value of the exchange rate at time t. Therefore, F_t is the prediction of the previous day and the period corresponds to the number of days in the window size used. In this project, I produced the mean using the same amount of days as the window size used in the LSTM implementation (30 days).

5.6 ARIMA - Autoregressive Integrated Moving Average

The Arima forecasting method is a fairly simple stochastic time series model that can be used to train and forecast future values of a time series. The acronym ARIMA stands for Autoregressive Integrated Moving Average and it is widely used for time-series predictions [12].

AUTOREGRESSION A regression model that uses the previous observations as input to predict the next one with an regression equation.

INTEGRATED Making the time series stationary so that we can integrate both models by measuring the differences of observations at different time [13].

MOVING AVERAGE Calculating values based on the values of the passed observations.

It works accurately under the assumption that the data inputted is a stationary univariate series, meaning that its mean and variance should not change over time [12]. In the context of this project, the forex exchange prices might vary in variance and mean, but its effect should be kept minimal since the data sets that we are using at are mostly constant. Regardless, I differentiated the series in order to make it stationary and prevent any changes in the mean or variance to prevent accurate ARIMA predictions. A specific model is often represented with ARIMA(p,d,q) where p is the number of lagged observations, d is the degree of differentiation and q is the window size of the moving average [13]. In the case of this project, I used the model ARIMA(2,1,0) since it has been found to be the most appropriate model for forecasting a financial indicator [14, p4].

5.7 Time Frame of Forecasts

We will forecast the models' ability to forecast values of the exchange rates for different time frame to see if the models are better at predicting the value of the next possible value or a relatively larger time frame.

1 DAY FORECAST Calculating values based on the values of the passed observations to predict the value of the next day. I suspect that a linear model like the ARIMA and the Exponential Moving Average might be better in this type of forecasting because of the linear dependency of the value to be predicted and the past values.

5 DAYS FORECAST Making a prediction based on the passed observations to predict the value of the exchange rate in 5 days. Therefore, there is a gap of unknown data between time T and the time that we want to predict (T+5). I suspect that the LSTM might be better in this time frame relative to the other models as it does not depends as strongly as the other models on linear relationships.

6 RESULTS

6.1 GRB/USD Data Set

6.1.1 *1 Day Predictions*

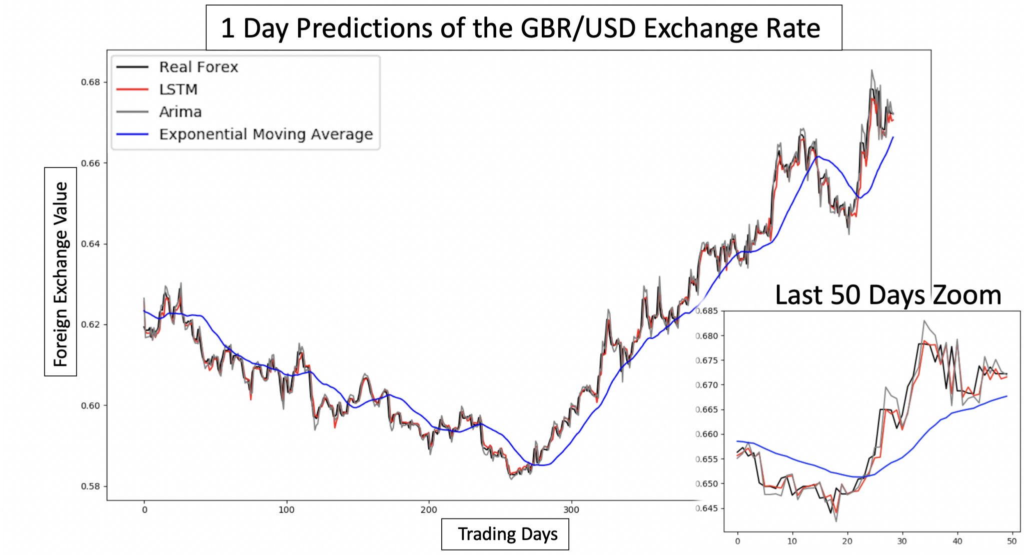


Figure 4: Results of the models with 150 epochs and a window size of 30 trading days for the 5 days forecast.

GBR/USD Rates 1 DAY FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.002486**	**0.001875**	**0.302027**
ARIMA	0.002491	0.001932	0.377102
Exponential Moving Average	0.005387	0.004215	1.097485

Table 2: Errors of the different models according to the different metrics on GBR/USD.

The best models of the 3 is the LSTM according to the 3 metrics. However, the ARIMA predictions are just a little bit less effective than the LSTM (only a difference of 0.000004 in the RMSE). The LSTM makes better prediction but the differences in the results might not be enough to affirm that one model is better than the other. Looking at the graph and interpreting the metrics, we can see that the LSTM errors were approximately in the same range for all of the predictions. A confidence interval will be calculated based on the T distribution in the next section to know with a 95%confidence level wheter one is better than the other. On the other hand, the exponential moving average's results are considerably worst result than the results other two models.

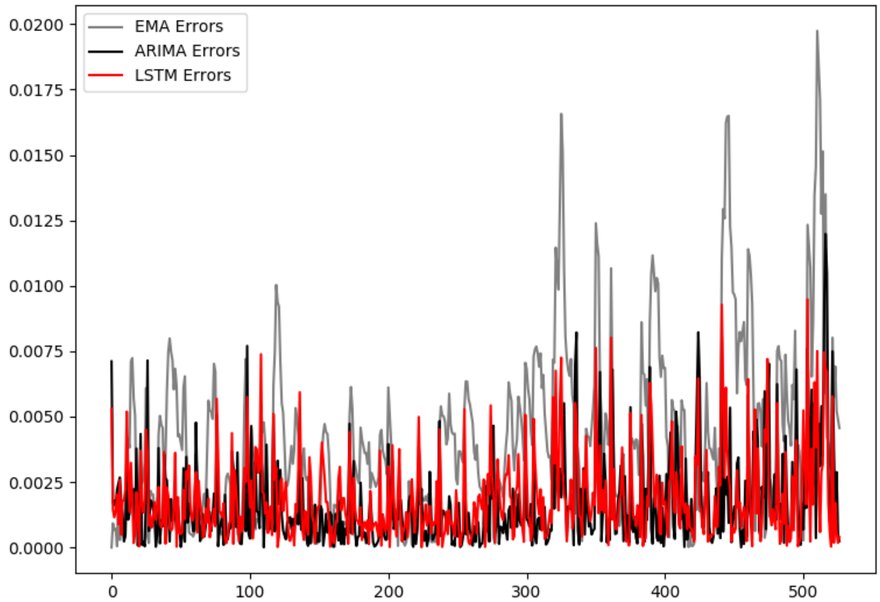


Figure 5: Absolute values of errors of the models calculated with the 600 trading days

6.1.2 *5 Days Predictions*

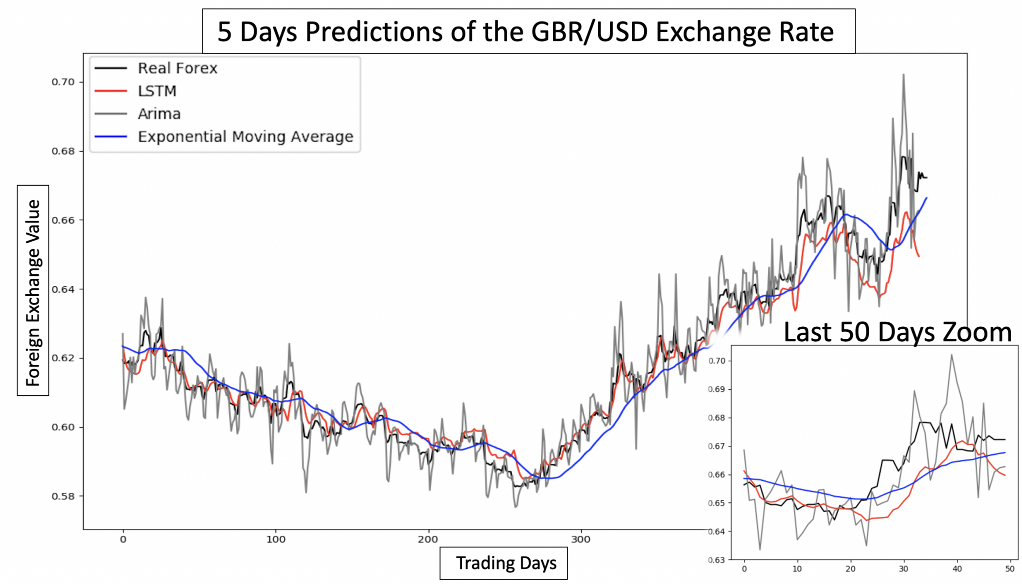


Figure 6: Results of the models with 150 epochs and a window size of 30 trading days for the 5 days forecast.

GBR/USD Rates 5 DAYS FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.006642**	**0.005054**	**0.803613**
ARIMA	0.007677	0.005962	4.422766
Exponential Moving Average	0.007076	0.005622	0.899738

Table 3: Errors of the different models according to the different metrics on GBR/USD.

For the 5 days forecasts of this data set, we see that the results of the LSTM (according to the 3 metrics) are considerably better than the other comparative models. Further analysis with the confidence interval will help in confirming this assumption.

6.2 USD/CAD Data Set

Results of the models for 1 DAY FORECASTS after 150 epochs on the USD/CAD data set.

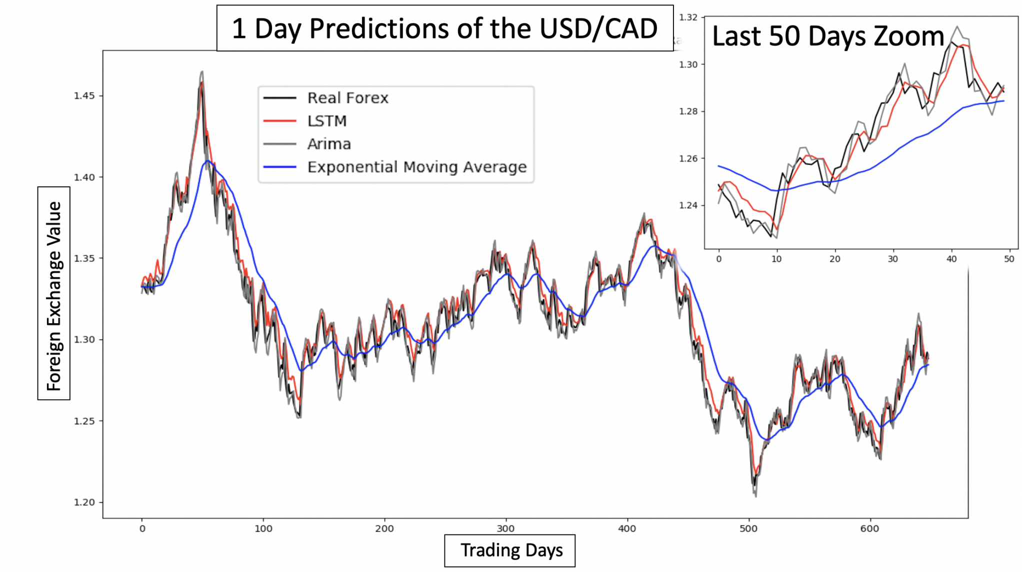


Figure 7: Results of the models with 150 epochs and a window size of 30 trading days.

USD/CAD Rates 1 DAY FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.006867**	**0.005267**	**0.401583**
ARIMA	0.00735	0.005599	0.427603
Exponential Moving Average	0.018089	0.01438	1.097485

Table 4: Errors of the different models according to the different metrics on USA/CAD.

The best of the 3 models presented is the LSTM according to the 3 metrics. The ARIMA predictions are relatively close to the ones of the LSTM, the MAPE metrics is only 0.0261 larger (which is the smallest gap of the 2 models between the three metrics). Since the difference between the results are considerable small, the confidence interval will help to determine if the LSTM is always better, or it was just a result of chance.

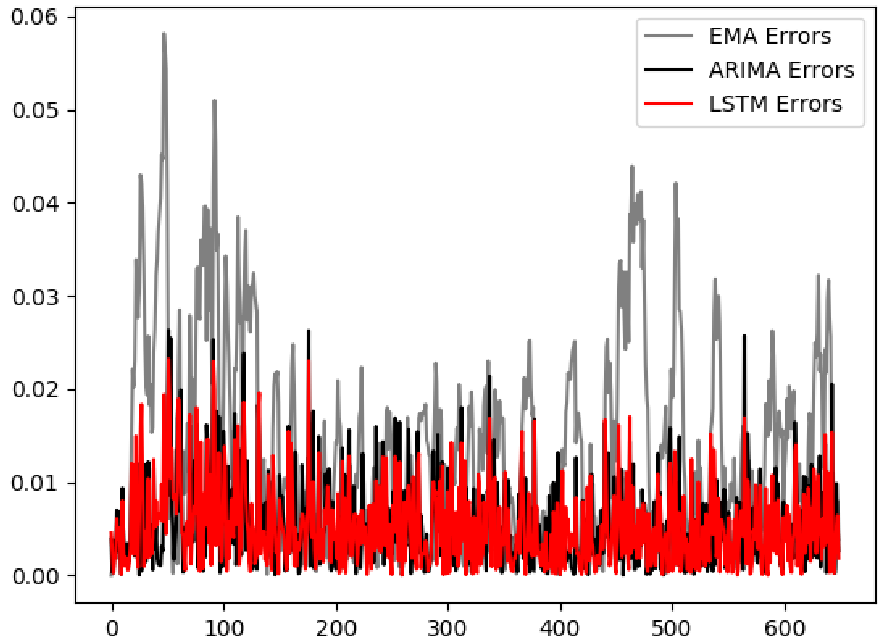


Figure 8: Absolute values of errors of the models calculated with the 600 trading days

Results of the models for 5 DAYS FORECASTS after 150 epochs on the USD/CAD data set.

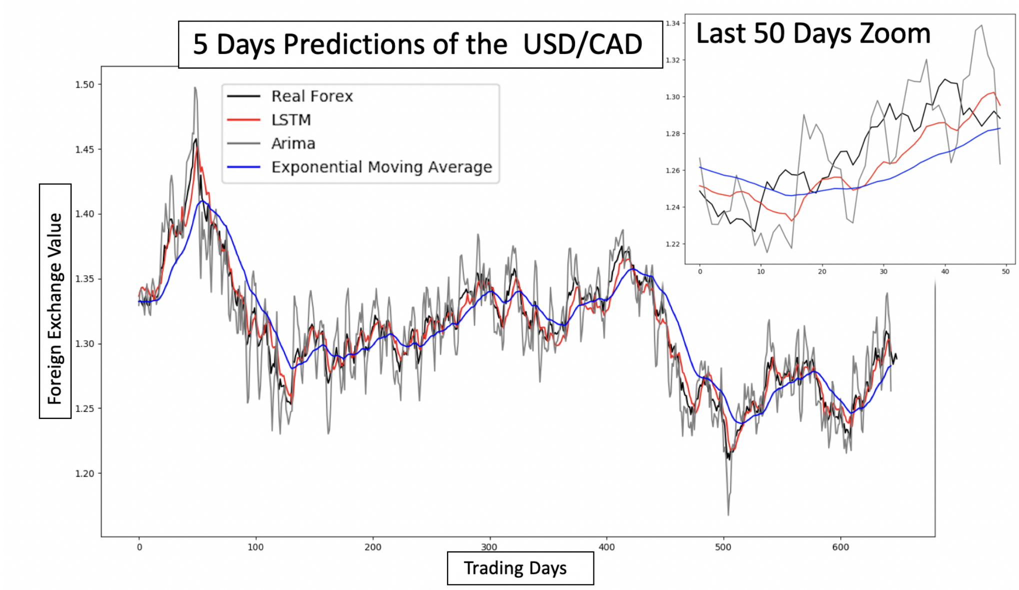


Figure 9: Results of the models with 150 epochs and a window size of 30 trading days.

USD/CAD Rates 5 DAYS FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.017631**	**0.014253**	**1.087542**
ARIMA	0.024187	0.019174	4.010076
Exponential Moving Average	0.023810	0.019034	1.452483

Table 5: Errors of the different models for the 5 days forecast according to the different metrics on GBR/USD.

For the 5 days forecasts of this data set, we see that the results of the LSTM (according to the 3 metrics) are considerably better than the other comparative models. Further analysis with the confidence interval will help in confirming this assumption.

6.3 AUS/USD Data Set

Results of the models after 150 epochs on the USD/CAD data set.

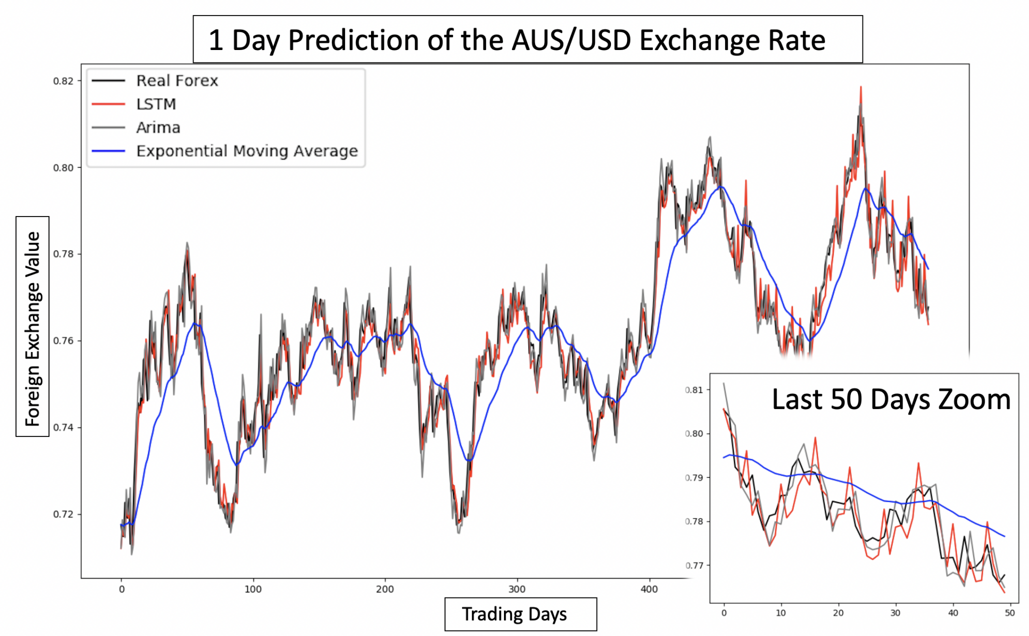


Figure 10: Results of the models with 150 epochs and a window size of 30 trading days.

AUS/USD Rates 1 DAY FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.005851**	**0.004591**	**0.356478**
ARIMA	0.006412	0.004885	0.379782
Exponential Moving Average	0.02138	0.017607	3.538286

Table 6: Errors of the different models according to the different metrics on AUS/CAD.

The LSTM model is better than the two other comparisons models. It is only slightly better than the ARIMA model, but it is the larger difference in the metrics between the 2 models for the 1 day forecasts. The confidence interval of 95% will determine if we can assume if it is strictly better than the ARIMA model.

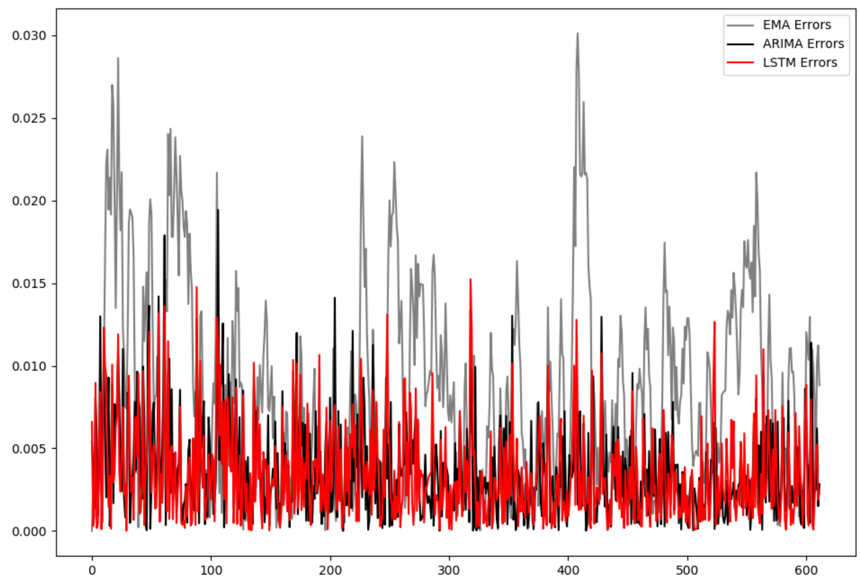


Figure 11: Absolute values of errors of the models calculated with the 600 trading days

Results of the models for 5 DAYS FORECASTS after 150 epochs on the AUS/USD data set.

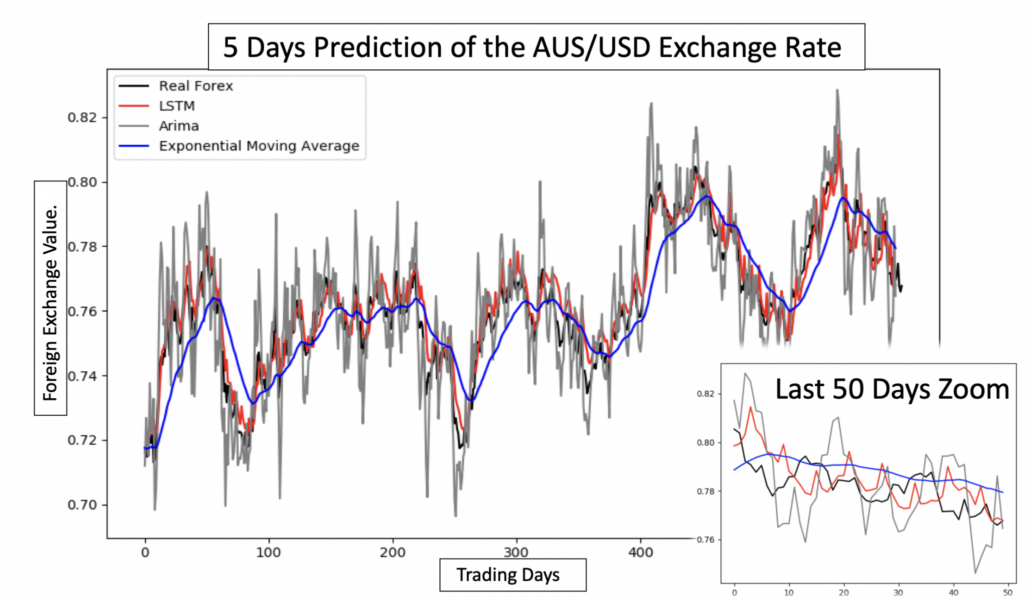


Figure 12: Results of the models with 150 epochs and a window size of 30 trading days.

AUS/USD Rates 5 DAYS FORECASTS			
	RMSE	MAE	MAPE
LSTM	**0.012894**	**0.010391**	**1.372681**
ARIMA	0.015011	0.011895	3.282825
Exponential Moving Average	0.013813	0.011334	1.488832

Table 7: Errors of the different models for the 5 days forecast according to the different metrics on AUS/USD.

For the 5 days forecasts of this data set, we see that the results of the LSTM (according to the 3 metrics) are considerably better than the other comparative models. Further analysis with the confidence interval will help in confirming this assumption.

6.4 Overall

The following 2 tables will present the results of one run of each of the data sets according to the time frame of the forecasts. A confidence interval is calculated in the next section to provide more accuracy in affirming that one is better than the others and are not due to chance.

Exchange Rates 1 Day Forecasts									
	GRB/USD			USD/CAD			AUS/CAD		
	RMSE	MAE	MAPE	RMSE	MAE	EMA	RMSE	MAE	EMA
LSTM	**0.00248**	**0.00188**	**0.30203**	**0.00687**	**0.00527**	**0.40158**	**0.00585**	**0.00459**	**0.35647**
ARIMA	0.00249	0.00193	0.37710	0.00735	0.00556	0.42760	0.00641	0.00488	0.37978
EMA	0.00539	0.00422	1.09749	0.01808	0.0143	1.09748	0.0213	0.01760	3.53828

Table 8: Results on the 3 different data sets according to the 3 comparison metrics for the 1 Day Forecasts.

Exchange Rates 5 Days Forecasts									
	GRB/USD			USD/CAD			AUS/CAD		
	RMSE	MAE	MAPE	RMSE	MAE	EMA	RMSE	MAE	EMA
LSTM	**0.00664**	**0.00505**	**0.80361**	**0.01763**	**0.01425**	**1.08754**	**0.01289**	**0.01039**	**1.37268**
ARIMA	0.00767	0.00596	4.42276	0.02418	0.01917	4.01008	0.01501	0.01190	3.28282
EMA	0.00708	0.00562	0.89974	0.02381	0.01903	1.45249	0.01381	0.01133	1.48883

Table 9: Results on the 3 different data sets according to the 3 comparison metrics for the 5 Days Forecasts.

We can see that the LSTM predictions are more precise with the three different data sets according to all of the error metrics. However, some of those metrics are extremely close, especially in the 1 Day Forecasts where the ARIMA have similar results to the LSTM. After running them more than 1 time, I realised that, in the 1 day forecasts, the LSTM was not always more precise (according to the metrics) than the ARIMA model. Therefore, in the next section, I will calculate a confidence interval for the 2 types of forecasts with the 3 data sets to determine with precision if the proposed model of LSTM is better than the others. In general, the errors of the LSTM are smaller than the Autoregressive Integrate Moving Average and the Exponential Moving Average but the confidence intervals will confirm if those results are a fact or are due to luck. In most cases of the 1 day predictions, the mean average error of the LSTM and the ARIMA model are really similar, ARIMA's results are slightly higher (approximately 6% higher). In the 5 days prediction, the LSTM seems to be the unambiguous choice as all the metrics seems to be pretty much smaller than the one's of the two other models, the confidence intervals will help determine if it is stricly better.

6.5 Verification of Assumptions with Confidence Intervals

In order to confirm if the LSTM model is better than the others, I ran the programs 10 times for the 2 time frames of the predictions for the 3 data sets. Afterwards, a confidence interval was determined with the T-Table distribution with a 95% confidence level that the value contains the true mean of the sample. The interval is considered with the RMSE metric as it is the indicator that I seen the most commonly in the comparison of predictive models.

1 Day Forecasts GBR/USD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$1.745E^{-4}$	$2.521E^{-3}$	$2.3726E^{-3}$	$2.6694E^{-3}$	$2.491E^{-3}$	$6.607E^{-3}$

Table 10: Confidence Interval Info For 1 Day GBR/USD

Taking into consideration the table above, we cannot confirm with a confidence level of 95% that the LSTM is better than the ARIMA since the result of the comparative model is inside of the interval of confidence. Further studies have to be conducted to determine if one is better than the other. However, we can confirm that it is better (according to the RMSE) than the Exponential Moving Average by a considerable margin.

1 Day Forecasts USD/CAD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$5.080E^{-4}$	$7.451E^{-3}$	$7.0199E^{-3}$	$7.883E^{-3}$	$7.350E^{-3}$	$1.809E^{-2}$

Table 11: Confidence Interval Info For 1 Day USD/CAD

Again, we cannot confirm that the LSTM is better than the ARIMA as it is inside of the confidence interval. However, We can say that it is better than the Exponential Moving Average (according to the RMSE). The ARIMA is likely better as its RMSE is lower than the mean of the LSTM predictions' RMSE. However, further analysis could help determine the best model between the ARIMA and the LSTM.

1 Day Forecasts AUS/USD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$2.462E^{-4}$	$4.568E^{-3}$	$4.359E^{-3}$	$4.777E^{-3}$	$4.782E^{-3}$	$1.054E^{-1}$

Table 12: Confidence Interval Info For 1 Day AUS/USD

From the results of the confidence interval of this data set, we can state that we are 95% confident that the LSTM is better (according to the RMSE) than the ARIMA model and the Exponential Moving Average. The confidence interval does not include the values calculated for the ARIMA and the Exponential Moving Average. It is the only data set in the 1 Day Forecasts that we can affirm we a 95% confidence level that the LSTM is a more precise model than the other two comparative models.

5 Days Forecasts GBR/USD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$6.032E^{-4}$	$6.580E^{-3}$	$6.069E^{-3}$	$7.093E^{-3}$	$7.677E^{-3}$	$7.284E^{-3}$

Table 13: Confidence Interval Info For 5 Days GBR/USD

It can be stated with a 95% level of confidence that the LSTM results are more precise (according to the RMSE) than the other two models since the results of those are higher than the upper bound of the confidence interval. An interesting fact to notice is that the results of the RMSE of the Exponential Moving Average is lower than the ARIMA, which was clearly not the case for all of the 1 day forecasts.

5 Days Forecasts USD/CAD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$1.166E^{-3}$	$1.860E^{-2}$	$1.761E^{-2}$	$1.959E^{-2}$	$2.419E^{-2}$	$2.381E^{-2}$

Table 14: Confidence Interval Info For 5 Days USD/CAD

According to this sample, we can also state that, with a 95% level of confidence, that the LSTM results are more precise (according to the RMSE) than the ARIMA and the Exponential Moving Average. Once again, even if they are pretty similar, the Exponential Moving Average result of the RMSE is smaller than the ARIMA.

5 Days Forecasts AUS/USD - RMSE					
STDEV	Mean	LowerBound	UpperBound	ARIMA	EMA
$8.680E^{-4}$	$1.162E^{-2}$	$1.088E^{-2}$	$1.236E^{-2}$	$1.501E^{-2}$	$1.381E^{-2}$

Table 15: Confidence Interval Info For 5 Days AUS/USD

As with the other 2 sets of 5 days predictions, we can affirm with a confidence level of 95% that the results of the proposed LSTM is better than the other two comparative models, according to the RMSE. We can confirm that because the RMSE results of the ARIMA and the EMA are higher than the upper bound of the confidence interval established. We can also witness that the Exponential Moving Average's RMSE is smaller than the ARIMA for the 3 data sets in the 5 days predictions.

7 DISCUSSION

In conclusion, this study has experimented with a few variations of LSTM but concluded that one including some major macroeconomic indicators of the concerned countries to predict the foreign exchange rates was of the most promising. The LSTM's architecture has been optimized and compared with comparative time series predictive models: the Autoregressive Integrated Moving Average and the Exponential Moving Average. The results are compared according to 3 different metrics: the Root-Mean-Squared error, the Mean-Average error and the Mean-Average-Percentage error. The 3 different foreign exchange rates chosen in this study is the GBR/USD, the USD/CAD and the AUS/USD. From the results demonstrated above, we can confirm with a 95% confidence level that the predictive model of the LSTM is more precise (according to RMSE) on all of the 5 days forecasts tested than the ARIMA and the Exponential Moving Average. However, the evidence have not been conclusive enough to affirm that they are better on the 1 day forecasts on all of the 3 data sets. The LSTM occasionally have better results than the 2 comparative models, but we can only state with a 95% confidence level that the results of 1 day forecasts are more precise for the AUS/USD data set. The LSTM is probably better than the other two models on the 5 days forecasts because it does not depends as strongly on linear dependencies of past values. Ultimately, it would be interesting to study the financial viability of such kind of model (taking into account transaction costs) and calculating which time frame of prediction would be optimal in the long-run. Furthermore, one could include more metrics on the past values, like upward and downward momentum. Depending on the pertinence of those indicators, the results of the LSTM could be improve, maybe to the point that it is more conclusively better the ARIMA model on the one day forecasts.

8 REFERENCES

[1] K. Amadeo, "3 Factors That Affect Exchange Rates," *The Balance*, 22-Apr-2019. [Online]. Available: https://www.thebalance.com/how-do-exchange-rates-work-3306084. [Accessed: 23-May-2019].

[2] N. Donges, "Recurrent Neural Networks and LSTM," *Towards Data Science*, 25-Feb-2018. [Online]. Available: https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5. [Accessed: 23-May-2019].

[3] N. Sinha, "Understanding LSTM and its quick implementation in keras for sentiment analysis.," *Towards Data Science*, 19-Feb-2018. [Online]. Available: https://towardsdatascience.com/understanding-lstm-and-its-quick-implementation-in-keras-for-sentiment-analysis-af410fd85b47. [Accessed: 23-May-2019].

[4] C. Olah, "Understanding LSTM Networks," *Colah's Website*. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/. [Accessed: 23-May-2019].

[5] N. K. Lioudis, "How do national interest rates affect a currency's value and exchange rate?," *Investopedia*, 12-Mar-2019. [Online]. Available: https://www.investopedia.com/ask/answers/040315/how-do-changes-national-interest-rates-affect-currencys-value-and-exchange-rate.asp. [Accessed: 06-Jun-2019].

[6] B. Beers, "What indicators are used in exchange rate forecasting?," *Investopedia*, 12-Mar-2019. [Online]. Available: https://www.investopedia.com/ask/answers/021715/what-economic-indicators-are-most-used-when-forecasting-exchange-rate.asp. [Accessed: 23-May-2019].

[7] W. Miyamoto, T. L. Nguyen, and V. Sheremirov, "The Effects of Government Spending on Real Exchange Rates: Evidence from Military Spending Panel Data," *SSRN Electronic Journal*, 2016.

[8] K. Eckhardt, "Choosing the right Hyperparameters for a simple LSTM using Keras," *Towards Data Science*, 29-Nov-2018. [Online]. Available: https://towardsdatascience.com/choosing-the-right-hyperparameters-for-a-simple-lstm-using-keras-f8e9ed76f046. [Accessed: 06-Jun-2019].

[9] J. Brownlee, "How to Diagnose Overfitting and Underfitting of LSTM Models," *Machine Learning Mastery*, 01-Sep-2017. [Online]. Available: https://machinelearningmastery.com/diagnose-overfitting-underfitting-lstm-models/. [Accessed: 24-May-2019].

[10] L. Posey, "Implementing Moving Averages in Python," *Towards Data Science*, 28-Mar-2019. [Online]. Available: https://towardsdatascience.com/implementing-moving-averages-in-python-1ad28e636f9d. [Accessed: 04-Jun-2019].

[11] A. Milton, "Trading with Simple, Exponential, and Weighted Moving Averages," *The Balance*, 31-Dec-2018. [Online]. Available: https://www.thebalance.com/simple-exponential-and-weighted-moving-averages-1031196. [Accessed: 11-Jun-2019].

[12] A. Singh, "Build High Performance Time Series Models using Auto ARIMA in Python and R," *Analytics Vidhya*, 07-May-2019. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/08/auto-arima-time-series-modeling-python-r/. [Accessed: 13-Jun-2019].

[13] S. Chatterjee, "ARIMA/SARIMA vs LSTM with Ensemble learning Insights for Time Series Data," *Towards Data Science*, 13-May-2019. [Online]. Available: https://towardsdatascience.com/arima-sarima-vs-lstm-with-ensemble-learning-insights-for-time-series-data-509a5d87f20a. [Accessed: 11-Jun-2019].

[14] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction," *Journal of Applied Mathematics*, vol. 2014, pp. 1–7, 2014.

9 LINK TO SOURCE CODE

Click here to see the GitHub Repository of the source code created for this project: https://github.com/francoisdavid/LSTM-NeuralNetwork-Forex