Machine Learning Approaches to Investigate Fundamentals That Impact the Trends in Foreign Exchange Rates

by
Farzeem Jiwani

A Major Research Paper
presented to Ryerson University (renaming to Toronto Metropolitan University in progress*) [1]
in partial fulfillment of the requirements for the degree of

Master of Science
in the Program of
Data Science and Analytics

Toronto, Ontario, Canada, 2022
© Farzeem Jiwani, 2022

[1]* In April 2022, the university announced the new name of Toronto Metropolitan University, which will be implemented in a phased approach.

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A MAJOR RESEARCH PAPER (MRP)

I hereby declare that I am the sole author of this Major Research Paper. This is a true copy of the MRP, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University, to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

Farzeem Jiwani

MACHINE LEARNING APPROACHES TO INVESTIGATE FUNDAMENTALS THAT IMPACT THE TRENDS IN FOREIGN EXCHANGE RATES

Abstract

The aim of this paper is to examine disparate economic factors that influence the US and the Canadian dollar exchange rate using machine learning techniques. Exchange rate forecasting is one of the attractive and challenging issues in international economics, as many uncertain factors come into play concurrently. Therefore, improving the accuracy of forecasting exchange rates and analyzing the fundamentals is of great interest to market participants, policymakers, and academics. One of the prominent issues for exchange rate prediction is the features driving it. The challenge of choosing the optimal forecasting model is mostly a result of the frequent changes among the factors impacting exchange rates. By taking into account a variety of alternative fundamentals that affect the Canadian dollar, we statistically evaluate the forecasting performance of the models and focus on understanding what drives movements in the Canadian dollar, and which indicators influence the CAD dollar the most. There are diverse attributes that could prospectively impact the exchange rate, e.g., the unemployment rate, purchasing power, oil price volatility and money supply. We forecast the exchange rate using traditional models, in particular, Random Forests, Extra Trees, XGBoost, Support Vector Machines, Lasso, and Ridge, as well as deep learning models, namely, LSTM and GRU. Evaluation is carried out using out-of-sample testing against standard regression metrics, specifically normalized deviation (ND), mean absolute error (MAE), root mean squared error (RMSE), and normalized root mean squared error (NRMSE). In addition to modeling the selected rates and indices mentioned above as potential features, the study aims to interpret the results using techniques such as SHAP to quantify the amount of significance each macroeconomic component attributes to change in the exchange rate. Our results indicate that the linear regression models namely, Lasso and Ridge regression, perform better than the others on spatial information. However, the deep learning models outperform others when trained on temporal information. While purchasing power parity, unemployment, money supply stocks, and oil prices were generally major factors in driving the exchange rate movement, the S&P 500 and commodity price indexes came out as significant for deep learning models. The need for a model that can predict foreign exchange rate direction and account for the relationship between these macroeconomic fundamentals and the exchange rate, will be an outstanding contribution as this helps understanding which economic factors drive the fluctuations in the exchange rates.

Acknowledgements

I am very grateful to Professor Mucahit Cevik and Professor Mohamed Wahab Mohamed Ismail for their support and assistance in making this project a reality. Professor Mohamed Wahab Mohamed Ismail was my supervisor for this MRP; he has been a great support throughout the term to guide and direct my research and provide valuable feedback. Professor Mucahit Cevik has been a motivation throughout the course to push me towards constant learning that has significantly helped me in this course.

Thank you, Professor Mucahit Cevik and Professor Mohamed Wahab Mohamed Ismail.

Contents

1 Introduction **5**
1.1 Research Objectives 6
1.2 Contributions 6
1.3 Organization of the Paper 7

2 Literature Review **7**
2.1 Review of Time Series Forecasting Methods 7
2.2 Time Series Forecasting in the Financial Domain 7
2.3 Comparison with Closely Related Works 8

3 Methodology **10**
3.1 Time Series Forecasting Basics 10
3.2 Dataset Description 11
3.3 Time Series Forecasting Models 13
3.4 AI Interpretability Methods 15
3.5 Experimental Setup 16
3.5.1 Model Training and Hyperparameter Tuning 16
3.5.2 Evaluation 17

4 Numerical Results **19**
4.1 Macroeconomic Features Only Models 19
4.2 Macroeconomic Features and Lag Values 19
4.3 Lag Values Only Models 20
4.4 Feature Importance Estimation 24

5 Conclusions **27**
5.1 Challenges and Study Limitations 27
5.2 Future Research Directions 30

1 Introduction

The foreign exchange market is one of the most important financial markets, due to its high liquidity levels. Given that theoretical and political considerations impact the foreign exchange rates, the research and forecasting of the money market's future values have received a great deal of interest. With the onset of the global financial crisis in 2008, academics accelerated their empirical work on financial crisis prediction and the creation of early warning systems. Economists and decision-makers became aware of the disastrous effects that detrimental commercial shocks might have on the actual economy. The crisis' crucial adverse consequence in this regard has been the sharp rise in the unemployment rate. As a result, this causes an impact on the global economy which makes this domain critical for legislators. The FOREX market is studied using fundamental analysis and technical analysis. Technical analysis only forecasts the FOREX market using historical time-series data, whereas fundamental analysis accounts for the company and the nation's economic and industrial conditions. In our paper, we focus on the US dollar to Canadian dollar foreign exchange rate prediction using technical analysis, whereas we derive the attributes based on fundamental analysis.

The Canadian dollar plays a significant role in the economic and political life of the nation. Although, it is often unclear how the Canadian dollar is affected or the impact that macroeconomic fundamentals and other financial factors play in projections for the Canadian dollar. Recently, the strength or weakness of the Canadian dollar against the US dollar has been discussed more often. The Canadian dollar has been depreciating since 2013. Governments and financial institutions are paying close attention to the behavior of the Canadian dollar. Most investors and government financial policymakers depend on the strength of the Canadian dollar. For example, when the Canadian dollar is weaker than the US dollar, US manufacturers tend to shift their business to Canada so they can cut costs [40, 7]. Likewise, policymakers also decide monetary policies, such as interest rates, based on the value of the Canadian dollar. Since the Canadian dollar can be freely traded on the open market, its value is determined by many fundamental factors that affect its supply and demand. Such factors increasing demand for the Canadian dollar put downward pressure on correlated exchange rates such as the US dollar and the British pound. Similarly, driven by fundamentals, an increase in the supply of the Canadian dollar will result in a lower exchange rate relative to other currencies. Hence, it is paramount to forecast the Canadian dollar based on sound macroeconomic and financial fundamentals to help governments and financial institutions make strategic decisions. The study answers potential questions about specific factors to focus on, that drive the dollar value significantly.

The majority of the existing work detailed in Section 2, is applied to either univariate time-series exchange rate forecasting that cannot capture the effect of other financial indexes on the exchange rate, or do not account for an exhaustive list of economic factors. This is a concern if we are interested in examining trends of the dollar value against movement patterns of these fundamentals. As a result, our key motivation of this study is to forecast the exchange rate in a multivariate manner as well as inspect the vital components in the finance domain that drive the USD-CAD dollar value.

1.1 Research Objectives

The lack of consistency and accuracy in foreign exchange rate prediction leads to several attempts seeking appropriate models, tools and techniques for predicting this market. Another challenge for this study is investigating the relationship between foreign exchange rate movement and some leading macroeconomic variables such as oil prices, gold, unemployment and key financial indexes.

The goal of this study is to forecast the USD-CAD exchange rate using novel machine learning techniques based on historical time-series data. 13 macroeconomic fundamentals were selected on subject matter expertise and previous literature studies. This study investigates the impact of different macroeconomic variables, namely, Oil Prices, Money Stocks, S&P 500 Index, and so forth, on the Canadian dollar exchange rate. One of the derived objectives of this study is to identify the causal relationship between these variables and identify the impact of their costs to the economy.

We have conducted a comparative analysis among various traditional models and deep learning networks to forecast one-step and multi-step exchange rates. For the comparative survey, we have chosen models from broadly three types namely, classical, tree-based and deep-learning models. In the literature review section, we have provided respective studies that support the selection of our models. The evaluation was done using in-sample results against standard regression metrics, namely, MAE, ND, RMSE, and NRMSE. Additionally, we interpret the best-performing models using a SHAP (SHapley Additive exPlanations) [28] technique, which is a game theoretic approach to explain the output of any machine learning model.

1.2 Contributions

The contributions of this paper can be summarized as follows:

- Our pipeline incorporates disparate financial variables from multiple sources to provide a holistic understanding of the impact of the exchange rate on the economy. The study comprehends the effect of indicators such as S&P 500 and TSX, that directly influences the global economy. Generally, the Money stocks, Oil prices, S&P 500 Index, Purchasing Power Parity, and Unemployment rate were found to be most influential factors.

- We propose a framework to interpret the models using a game-theory method to investigate the magnitude of influence the financial attributes have on the USD-CAD exchange rate. These results can potentially help the government representatives and policymakers formulate judicious guidelines for the betterment of the economy.

- We have performed a comparative survey of traditional and deep learning techniques. The models are selected based on previous literature elaborated in Section 2 of the paper. These models are used to forecast one-step and multi-step predictions (5-step and 10-step, in particular), respectively. Linear Regression models had better performance when the features were used, however RNN models achieved predominant results when temporal information was incorporated.

1.3 Organization of the Paper

The remainder of the paper is organized as follows. Section 2 provides an overview of the relevant studies on time series forecasting in the FOREX domain and a brief on their respective implementations. Section 3 introduces our Methodology, outlining the steps performed in the project. Section 4 details the experimental setup in terms of evaluation techniques, metrics, and model parameters, followed by numerical results. Lastly, Section 5 concludes the paper with a summary of our findings and a discussion on future research directions.

2 Literature Review

Below, we first provide a literature review of the time-series forecasting methods. Then, we review the time-series forecasting methodologies in the financial domain. Lastly, we review the most relevant studies and position our research among the current literature.

2.1 Review of Time Series Forecasting Methods

Time series forecasting has been a prominent research field with applications in various domains and it has undergone major methodological advancements over recent years. Athiyarath et al. [1] covered and compared various forecasting algorithmic approaches and explored their limitations and usefulness for different types of time series data in different domains. Earlier studies such as Box et al. [4], focused on linear statistical models such as auto-regressive, moving average (MA) and auto-regressive integrated moving average (ARIMA), which account for linear correlations between past data points to make future predictions. With growing availability of exogenous variables, ML models such as Random Forests (RF), Support Vector Machines (SVM) and Extreme Gradient Boosting (XGB) grew in popularity for their effectiveness in dealing with cross-sectional feature spaces, as shown in Liu et al. [27].

More recently, RNNs such as Long Short Term Memory (LSTM) [18] and Gated Recurrent Unit (GRU) [10] architectures have been frequently employed for forecasting tasks due to their ability to extract long-term dependencies between temporal sequences. These networks can not only process non-linear data, but also retain memory for the sequence and retain useful information, which is positive.

2.2 Time Series Forecasting in the Financial Domain

Forecasting research literature is rich in terms of the published work in recent times owing to the development of information technology. In this sub-section, we look at prior time-series forecasting work in the financial domain. Karanikola et al. [21] performed a comparative review of various contemporary methods namely, ARIMA, Neural Basis Expansion Analysis (NBEATS), and Probabilistic Time Series Modeling on 40 univariate time series of financial data. Likewise, Yang and Wang [43] compared the performance of

ARIMA, support vector regression (SVR), and bidirectional long short-term memory neural network (BiLSTM) to predict the long-term and short-term dynamic trends of financial time series effectively. Cochrane [11] covered different techniques for applying time-series forecasting to macroeconomics and the financial domain. Krollner et al. [25] surveyed recent literature in the domain of machine learning techniques and artificial intelligence to provide researchers with a cohesive overview of recent developments in stock market index forecasting.

The experimental results show that as opposed to the statistical models such as ARIMA, neural network models produce better results, demonstrating their suitability for forecasting the foreign exchange rates. Guo [17] used 5000 observations from the S&P 500 index for empirical research, and introduced benchmark models, such as ARIMA, and GARCH for comparison, to verify the effectiveness and advantages of LSTM models. Khandelwal et al. [22] expanded it's study to compare Convolution Neural Network (CNN) along with ARIMA, BiLSTM, and unidirectional LSTM on financial time-series data. Ghadimpour and Ebrahimi [16] provided further evidence of deep learning techniques by comparing LSTM and GRU, in forecasting daily movements of the Standard & Poor (S&P 500) index using the daily closing price of this index. Park and Yang [32] presented a deep learning model based on the LSTM network architecture to predict economic growth rates and crises by capturing sequential dependencies within the economic cycle. In addition, the study provided an interpretable machine learning model that derives economic patterns of growth and crisis through efficient use of the eXplainable AI (XAI) framework.

2.3 Comparison with Closely Related Works

We started reviewing existing studies for forecasting the FOREX rates. Sezer et al. [37] provided a systematic review of the different techniques for financial time series forecasting. They categorized the studies not only by their targeted forecasting implementation areas, such as index, currency, and commodity forecasting, but also by the Deep Learning models used, such as CNNs, Deep Belief Networks (DBNs), and LSTM. Aydin and Cavdar [2] compared the performances of neural networks and Vector Autoregressive (VAR) models by using Gold Prices, and Borsa Istanbul 100 Index (BIST) macroeconomic variables for predicting USD-Turkish Lira exchange rates, concluding the superior performance of ANN over VAR models. Maneejuk and Srichaikul [29] gave further evidence than Recurrent Neural Networks (RNN) and Support Vector Machines (SVM) achieved notably higher performance than the other traditional models, in terms of standard regression evaluation metrics such as MAE, MAPE, RMSE, and Theil U. Liao et al. [26], which forecasted the USD-RMB rate, is a testament to the power of RNN and SVM.

Apart from the strategies mentioned above, we discovered some alternative methods that can be used. Shen and Liang [39] implemented an Autoencoder model with SVM to forecast varying exchange rates and found that it performed better over the existing benchmarks on the same dataset. Conversely, Sarangi et al. [36] used a hybrid Artificial Neural Network and Genetic Algorithm (ANN-GA) approach to analyze the INR-USD forex trends. Panda et al. [31] showed that CNN had the best accuracy compared to traditional/simpler models like ARIMA, Linear Regression, and Multi-Layer Perceptron networks. Islam et al. [20] employed machine learning approaches, such as XGB, to create forecasting apps and analyze the USD/Rupiah exchange rate based on time series data. And lastly, Islam and Hossain [19] deployed a GRU-LSTM network for forecasting four major currency pairs: EUR/USD, GBP/USD, USD/CAD, and USD/CHF.

We then started looking at research papers that have aimed to study how the FOREX market can be potentially affected by different factors. Reboredo et al. [33] carried out a detrended correlation approach to identify how Oil and the US Dollar rates are dependent on each other. The study provided evidence of both contagion and interdependence during certain periods. Tang et al. [41] also focused on the role of Oil futures information in predicting the US market volatility by conducting multivariate analysis using the heterogeneous autoregressive (HAR) framework. Bruneau and Moran [5] portrayed how the Producer Prices are sensitive to these Exchange Rates. Dick et al. [12] demonstrated that FOREX forecasts are related to a proper understanding of fundamentals such as Purchasing Power Parity (PPP), Short Term Interest Rates (IR), and Commodity Price Index (CPI) of Canada. The above references discuss Oil as a likely factor but we decided to not limit the scope and instead find other macroeconomic attributes.

Economic theories and empirical evidence have identified several fundamentals that have predictable effects on exchange rates. The popular ones include macroeconomic indicators, such as purchasing power parity, interest rate parity, and monetary supply [3, 6, 9, 12, 15, 30, 35, 42]. As for the fundamentals affecting the Canadian dollar, several have been addressed in the literature. Kouwenberg et al. [24] used oil price as a fundamental for forecasting ten exchange rates including the Canadian dollar. Cao et al. [8] found that domestic and export prices are sensitive to the movement of the Canadian dollar. For example, the Canadian dollar positively correlates with the crude oil price. Since Canada is an oil exporter, the Canadian dollar tends to appreciate against other major currencies, including the US dollar, when crude oil price increases. Furthermore, the Canadian dollar is affected by global liquidity, which can be captured using the TED Spread. In addition, currency such as the Canadian dollar has a stronger correlation with commodity prices. A new study also examinds the relationship between unemployment and exchange rate fluctuations [5, 14]. However, TSX is the most critical stock index in Canada and connects highly with CAD, and it is rarely a fundamental. Hence, we explore a set of fundamentals impacting the Canadian dollar besides the key fundamentals used in the literature. That includes TSX, S&P 500, TED spread, oil price, and gold price. Potential fundamentals selected in this research are listed below. We consider industrial production as a proxy for GDP. Regarding money supply, we have the empirical study with M1 and M3 stocks separately. TED-spread is the spread between the 3-month interbank rate and the 3-month treasury yield and acts as a proxy of global liquidity. For forecasting the daily USD-CAD exchange rate, the final set of variables and the fundamentals selected in this study are outlined in Table 1.

The above studies helped us understand the diverse macroeconomic fundamentals to consider in constructing the final dataset and finally establishing the algorithms to use in the project. Based on this literature survey, we have decided to use Linear Regression (with L1/L2 Regularization), Random Forest, Extra Trees, SVR, XGBoost, GRU, and LSTM models for predicting the USD-CAD FOREX rates.

3 Methodology

In this chapter, we discuss the various strategies used for time series forecasting. We first provide a brief introduction on the basics of Time-Series forecasting, followed by literature about our fundamentals. Then, we describe our dataset, including the distribution of the features and general data characteristics. Then, we elaborate on the models and architectures employed in our analysis and assess their strengths and drawbacks. Finally, we take a detailed look at the proposed pipeline of our study.

3.1 Time Series Forecasting Basics

The time series dataset can be broken into three components namely, trend, seasonality, and noise. The trend is the upward and downward shift of the data points at every time step over a period, seasonality is the seasonal variance that occurs due to a recurring event, and noise is the randomness that occurs in the form of spikes and troughs at random intervals within the series. Moreover, a time series is considered stationary if, over some time, it does not show a continuous increase or decrease, and it is classified as non-stationary otherwise.

Forecasting is commonly for generating one-step or multi-step ahead predictions. For one-step-ahead predictions, a base model is trained such that the lag values up to time T, together with any exogenous features, are used to predict the value for time $T + 1$. For multi-step ahead forecasting, commonly, two strategies are used. The first strategy is the Direct Multi-step Forecast Strategy, which uses a base model to forecast for every time step in a time series. For example, in a scenario that requires n-step ahead predictions, a different base model is trained for each of those nth-step predictions. The advantage of this method is that, since it uses lags on the same time instance in a particular data series, it is simpler to implement and experiment. However, a downside to this approach would be the high computational cost of training each of the separate base models. The second approach is to use a multi-output prediction model which is capable of generating multiple predictions. These models can learn the dependencies between inputs and outputs in addition to those between model outputs.

Testing for time series data involves following specific protocols. Unlike many other AI applications, the data cannot simply be split randomly for training and validation, as doing so would disrupt the time component of the series. Instead, the data is split such that one portion of the data (from $T = 0$ to $T = k$) is set aside for training, and the remaining portion ($T = k + 1$ onwards) is for validation. Once the data is split appropriately, different strategies are used for testing. One technique would be to divide the test data into discrete batches of size N, where none of the batches overlap with each other. This testing approach is easy to implement and works best when the testing data is abundantly available. Another technique would be to create test data using a rolling window approach. This approach requires dividing the data such that each subsequent batch gets shifted ahead by a timestamp, such that batch b contains data points from $T = k$ to $T = k + N$, and batch $b + 1$ holds data points from $T = k + 1$ to $T = k + N + 1$. Since this approach allows overlapping between batches, more testing instances can be generated using a limited dataset. Lastly, the augmented-out-of-sample comparison method can be employed for training and testing, which updates the model after each test to incorporate the test data into the training set. Since the model is updated only on a small portion of the dataset, there is no need to retrain the entire model, which speeds up the testing. It has been demonstrated that this method can compare several models accurately and with a high degree of certainty.

Table 1: Data sources

Predictor	Description	Frequency	Source
USD-CAD Rate	Foreign Exchange Rate	Daily	Bank of Canada
IR	Short-term Interest Rate Parity	Monthly	OECD
PPP	Purchasing Power Parity	Yearly	OECD
PPI	Producer Price Indices Parity	Monthly	OECD
Money Supply	Money stock (M1, M3)	Monthly	OECD
S&P 500	Standard & Poor's Index	Daily	Economic Research
TSX	S&P/TSX Composite Index	Daily	Yahoo Finance
Gold	Gold Fixing Price 10:30 A.M. in London Bullion Market, based in U.S. Dollars.	Daily	Economic Research
Oil	WTI Spot Price FOB (Dollars per Barrel)	Monthly	Economic Research
ED Spread	3-Month London Interbank Offered Rate (LIBOR)	Daily	Economic Research
Unemployment	Unemployment Rate	Monthly	Economic Research
Commodity Price	Commodity Price Index	Monthly	Bank of Canada
Industrial Production	Industrial Production Index	Monthly	OECD

3.2 Dataset Description

The dataset consists of 13 fundamental attributes and the USD-CAD FOREX rate as the target variable. All the 13 attributes are numeric, so we have used charts to depict the trend of the same depending on their respective data frequency. The dataset period is from January 2009 to December 2021.

Datasets for this project are open-sourced, compliant with the MRP requirements, and collected from the sources mentioned in Table 1. Table 2 outlines the descriptive statistics of the dataset. Figure 1 illustrates the general trend of the USD-CAD foreign exchange rate from 2009 till the end of 2021. The graph is fairly non-linear as it decreases till the year 2012 but later increases with a spike in the years 2016-2017, and the years 2020-2021 (possibly due to the pandemic). Figures 2 and 3 depict the nature of the 13 attributes used in our dataset. Gold prices show an upward trend with a dip in 2015, and then increasing till the end of 2021. Oil prices show a similar pattern but drop in 2020, potentially due to the pandemic. The stock prices represent a fairly increasing trend with some decreasing spikes in the years 2020-2021, again, potentially due to the economic impact of the pandemic. Industrial production showing a remarkable decrease in 2020, while unemployment depicted a sharp rise in the same year. While purchasing power, interest rate, ED spread and commodity price index illustrate a non-linear pattern, the money supply stocks and producer prices depict an increasing linear trend.

Table 2: Summary statistics

	count	mean	std	min	25%	50%	75%	max
DEXCAUS	3253	1.19	0.14	0.94	1.04	1.24	1.31	1.46
SP500	3253	2155.15	922.89	676.53	1351.77	2049.58	2747.71	4793.06
TSX	3253	14339.02	2486.57	7566.90	12456.50	14376.20	15807.20	21768.50
Gold	3253	1385.19	256.35	813.00	1217.70	1311.10	1598.25	2061.50
ED	3253	0.77	0.75	0.11	0.26	0.36	1.15	2.82
BCPI	3253	501.54	111.34	243.70	410.36	480.07	613.72	721.83
IR	3253	1.03	0.50	0.18	0.74	1.16	1.18	2.19
PPI	3253	101.34	7.90	88.45	97.84	100.08	105.84	126.01
M1	3253	107.62	37.48	57.68	77.62	99.24	125.19	204.29
M3	3253	104.00	28.14	66.24	79.49	98.63	122.46	163.19
Oil	3253	68.80	21.84	16.55	49.82	67.73	89.04	110.04
Unemployment	3253	7.28	1.23	5.40	6.70	7.10	7.70	13.40
IndProd	3253	98.10	5.26	74.64	95.68	99.05	101.22	107.11
PPP	3253	1.23	0.02	1.20	1.21	1.23	1.25	1.29

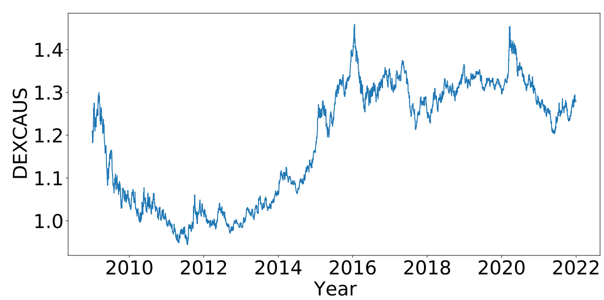
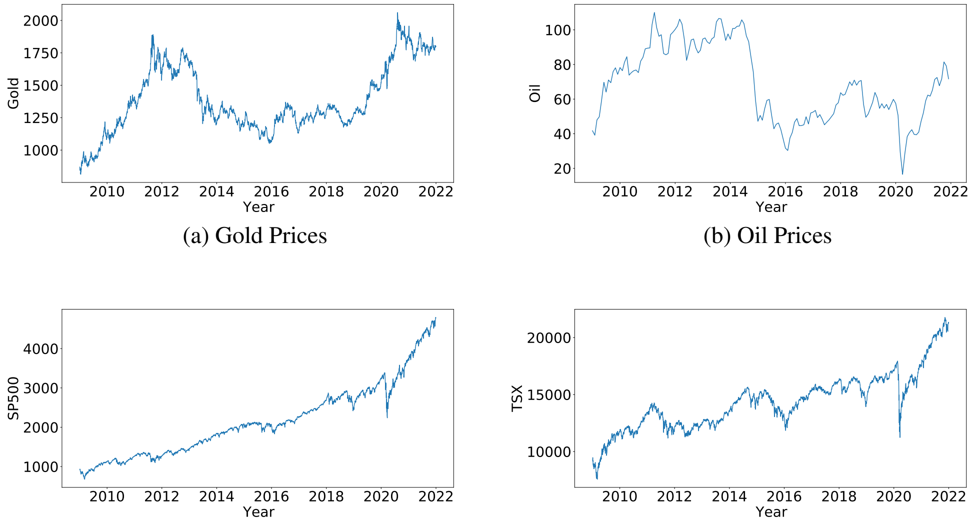


Figure 1: USD-CAD Exchange Rate Trend

Figure 4 represents the correlation matrix of the dataset. As expected, we see a high correlation between the stock market entities such as S&P500 and TSX with the Target variable. Moreover, we also see a notable high negative correlation between the exchange rate and oil prices and a moderately high positive correlation with the Money Supply stocks (M1/ M3).



(a) Gold Prices
(b) Oil Prices
(c) S&P 500
(d) TSX

Figure 2: Features for the Model Training: Commodities and Stock Prices

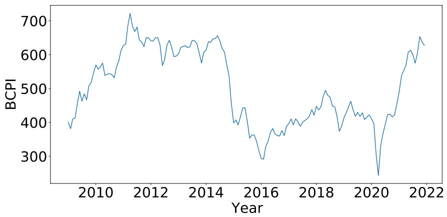
3.3 Time Series Forecasting Models

While a variety of machine learning models have been employed in the past to perform time series forecasting jobs, in our study, we use linear regression models, namely, Lasso and Ridge, and tree ensembles, in particular, Random Forests, ExtraTress, and XGBoost models. Note that we choose these models as they reportedly show high performance for various time series forecasting tasks.

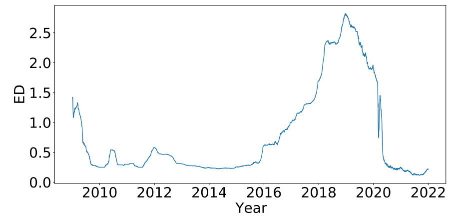
ML models require careful feature extraction for training high-performance forecasting models. Commonly extracted features for time series forecasting include features obtained from timestamps, e.g., time of the day, day of the week, and month of the year. For these features, one may additionally account one-hot encoding or sin/cosine transformation. The 13 economic factors, as explained in the dataset description section, make up our feature space, and we have limited our modeling to just these variables to comprehend their effects.

Recurrent Neural Networks process given information incrementally while maintaining an internal model of what is processed based on the past data and constantly updating its state as new information is received. They learn temporal dynamics by mapping an input sequence to a hidden state sequence and outputs via a recurrent layer and a feedforward layer. As such, they are suitable for problems where the sequence of the data matters.

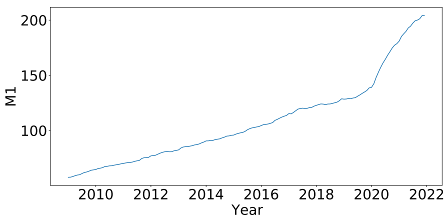
In our study, we have implemented Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) RNNs instead of Vanilla RNN to overcome the exploding and vanishing gradient problem. These networks have proven to yield exceptional results in prior research based on financial time-series forecasting use cases.



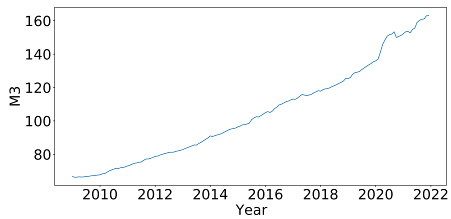
(a) BCPI



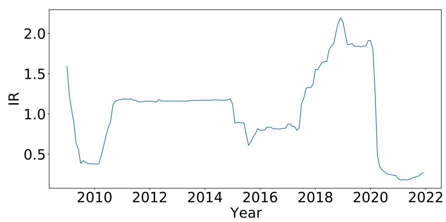
(b) TED Spread



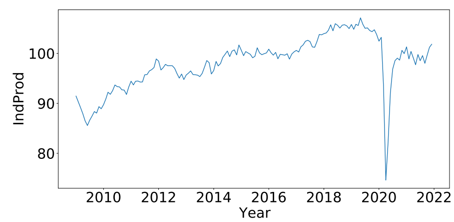
(c) M1 Money Supply



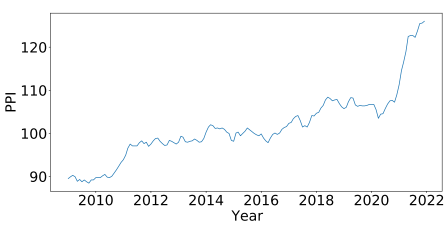
(d) M3 Money Supply



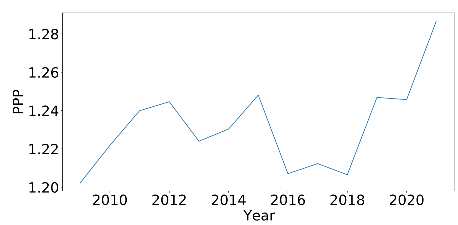
(e) IR



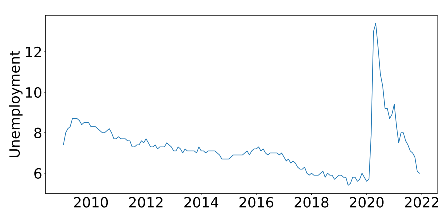
(f) Industrial Production



(g) PPI



(h) Purchasing Power



(i) Unemployment

Figure 3: Features for the Model Training: Macroeconomic Indicators

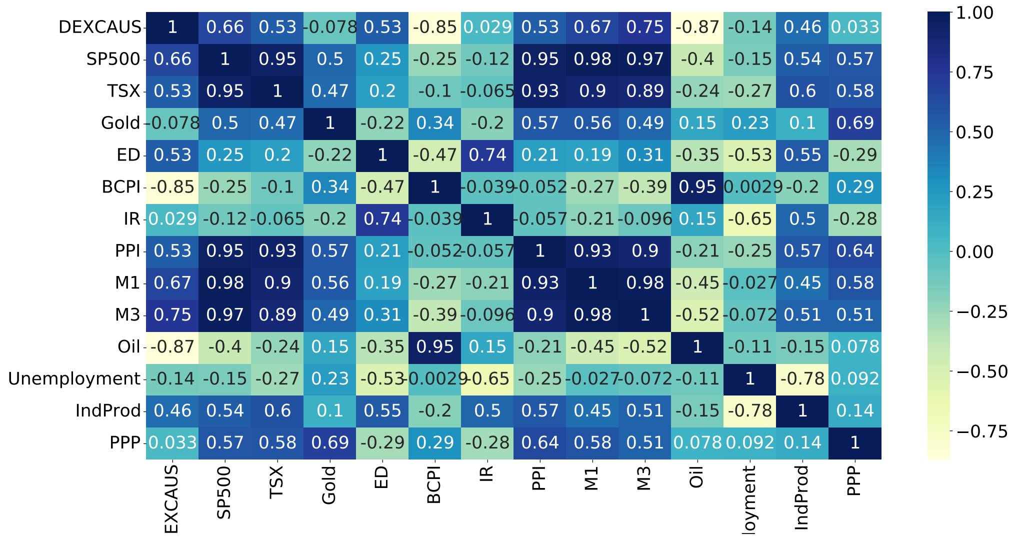


Figure 4: Correlation Matrix

3.4 AI Interpretability Methods

The level of interpretability refers to how well a person can consistently predict the outcome of the model [23]. If the decisions of one model are simpler for a human to understand than the decisions of the other model, then that model is considered more interpretable than the other model. The need for interpretability arises from an incompleteness in problem formalization [13], which means that for certain problems or tasks it is not enough to get the prediction (the what). Since a successful prediction only addresses a portion of your initial issue, the model must also provide a justification for how it arrived at the conclusion (the why). An explanation typically establishes a clear connection between the instance's feature values and the model's prediction. Robnik-Šikonja and Bohanec [34] details the characteristics of explanation methods and explanations, in order to assess how effective an explanation method or explanation is. To understand what are the main features that affect the model output, we need Explainable Machine Learning techniques that unravel some of these aspects.

In our study, we have used Shapley values, a method from coalitional game theory that tells us how to fairly distribute the payout among the features. The Shapley value, coined by Shapley (1953) [38], is a method for assigning payouts to players depending on their contribution to the total payout. Players cooperate in a coalition and receive a certain profit from this cooperation. The fact that the difference between the prediction and the average prediction is fairly spread throughout the feature values is one of this technique's main advantages.

SHAP [28] is a technique based on Shapley values, which explains how each feature affects the model and allows local and global analysis of the dataset and problem at hand. The goal of SHAP is to explain the prediction of an instance x by computing the contribution of each feature to the prediction. The feature values of a data instance act as players in a coalition. Shapley values tell us how to fairly distribute the payout (prediction) among the features. Linear models can use their coefficients as a metric for the overall importance of each feature, but they are scaled with the scale of the variable itself, which might lead to distortions and misinterpretations. Also, the coefficient cannot account for the local importance of the feature, and how it changes with lower or higher values. The same can be said for the feature importance of tree-based models, and why SHAP is useful for the interpretability of models. The idea behind SHAP feature importance is that the features with large absolute Shapley values are important. The average of the absolute Shapley values per feature across the data is plotted in the order of decreasing importance.

3.5 Experimental Setup

We conduct the numerical experiments using Scikit-learn version 1.0.2 and TensorFlow version 2.8.2 on a 2.7 GHz dual-core i5 processor with 8GB of RAM. All the implementations are in the Python programming language.

As covered in the Dataset description chapter, there are 13 economic indicators that we have used in our research. These variables are collected from their primary data sources based on their periodicity. The periods are either daily, monthly, quarterly or yearly. Subsequently, we consolidate these attributes using a repeating technique, i.e., we replicate the data to achieve a universal daily frequency since the target variable is being recorded daily. The dataset is divided into training and test sets using a data split of 85%. The training set consists of data from January 2009 till the end of December 2019, whereas the test set contains data points from January 2020 till the end of December 2021.

3.5.1 Model Training and Hyperparameter Tuning

This stage consists of two sets of models, traditional methods and deep learning networks. The conventional models comprise the Lasso, Ridge, SVM, Random Forest, Extra Trees, and XGBoost regression models. On the other hand, the deep learning method contains LSTM, GRU, and a hybrid LSTM-GRU network.

The modeling comprises three types of inputs, one with just the features and the other with the features and lag values, to understand the impact of lag values. Furthermore, we model the RNNs on the lag values solely to examine if it improves the predictive performance or not.

Since the data is only available for business days, we used a lag value of 5 days in our analysis. We perform one-step and multi-step predictions for both sets of models. For classical models, we integrate the results over multiple steps and note the evaluation metric's average value. The dataset for the deep learning models is reshaped based on the n-step window we wish to forecast and the lookback period. After all the multi-step predictions have been assessed, the mean value is reported. For 5-step predictions, we used a lookback period of 15 days. On the contrary, a lookback window of 30 days was used, for 10-step predictions.

Our experiments use random seed and random state values to mitigate the stochasticity involved with ML model training. We perform extensive hyperparameter tuning for all the forecasting models, where, for our deep architectures, we experiment with different combinations of stacked layers and hidden units, along with other parameters, including optimizer and batch size. For other ML models, we apply grid search to identify the best-performing parameters. The final set of hyperparameters used for each model is provided in Table 3, while the search space that compromises all different parameter combinations used in our hyperparameter tuning experiments is in Table 4.

Table 3: Hyperparameter settings used in the experiments for the employed models

Model	**Final Parameters**
Lasso	*alpha=0.01*
Ridge	*alpha=3000*
SVR	*C: 0.1, epsilon: 0.1, gamma: 1, kernel: 'linear'*
Random Forest	*n_estimators=400, max_depth=9, max_features='sqrt'*
Extra Trees	*n_estimators=300, max_depth=8, max_features='sqrt'*
XGBoost	*n_estimators=1000, max_depth=3, colsample_bytree=0.5, learning_rate=0.1*
LSTM	*hidden units= {500, 250, 125}, activation = {'relu'},* *optimizer= Adam, loss= mse, batch size= 256, buffer size = 150,* *lookback (5 step)= 15, lookback(10 step) = 30*
GRU	*hidden units= {1000, 500, 250, 125}, activation = {'relu'},* *optimizer= Adam, loss= mse, batch size= 256, buffer size = 150,* *lookback (5 step)= 15, lookback(10 step) = 30*
GRU-LSTM	*hidden units= {512, 256}, activation = {'relu'},* *optimizer= Adam, loss= mse, batch size= 256, buffer size = 150,* *lookback (5 step)= 15, lookback(10 step) = 30*

3.5.2 Evaluation

We consider four performance evaluation metrics, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Normalized Deviation (ND), and Normalized Root Mean Squared Error (NRMSE), to compare the performances of forecasting methods. The formulas are illustrated in (1), where y_i denotes the actual values and $\hat{y}_i$ denotes the predicted values.

The mean absolute error (MAE) refers to the magnitude of difference between the prediction of observation and the true value of that observation. NRMSE is a popular metric for assessing the performance of regression models, and it is typically the preferred method when the model errors follow a Gaussian distribution. ND, likewise, increases linearly with an increase in deviations from the ground truth. Each batch in the test set is subject to these metrics individually, and the average across all the clusters is reported as the final performance value.

Table 4: Hyperparameter search space for forecasting models

Panel A: Deep Learning Models	
Parameter	Search Space
hidden layers	[1, 2, 3, 4]
hidden units	[32, 64, 125, 128, , 250, 256, 500, 512, 1000, 1024]
optimizers	[Adam, Adamax, SGD]
batch size	[32, 64, 128, 256]
Panel B: Classical Models	
Parameter	Search Space
Lasso: alpha	[0.0001, 0.001, 0.01, 0.1, 1]
Ridge: alpha	[100, 1000, 2000, 3000, 4000, 5000, 10000]
SVR: C	[0.1, 1, 1.5, 2, 5, 10, 100]
SVR: gamma	[1, 0.1, 0.01, 0.001, 0.0001, 0.00001]
SVR: kernel	['rbf', 'linear', 'poly']
SVR: epsilon	[0.1,0.2,0.3,0.4,0.5]
Panel C: Tree-Based Models	
Parameter	Search Space
n_estimators	[100, 300, 500, 700, 1000, 2000]
max_features	['sqrt', 'log2']
max_depth	[3,4,5,6,7,8,9,10]
learning_rate	[0.01, 0.05, 0.1, 0.15]
colsample_bytree	[0.3, 0.5, 0.7]

$$\text{MAE (y, ŷ)} = \frac{\sum_{i=1}^{N} |y_i - \hat{y}_i|}{N} \tag{1}$$

$$\text{RMSE (y, ŷ)} = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}} \tag{2}$$

$$\text{ND (y, ŷ)} = \frac{\sum_{i=1}^{N} |y_i - \hat{y}_i|}{\sum_{i=1}^{N} y_i} \tag{3}$$

$$\text{NRMSE (y, ŷ)} = \frac{\sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}}}{\frac{\sum_{i=1}^{N} y_i}{N}} \tag{4}$$

4 Numerical Results

In this section, we provide results from our numerical study and discuss our findings for the forecasting task. Without including the lag values, we first contrast the effectiveness of conventional and deep learning models. Then, we examine the effects of incorporating sequential data in deep learning models. Finally, we list the importance of each attribute for the top-performing models.

4.1 Macroeconomic Features Only Models

In this section, we report the evaluation metrics when the input to the model is just the 13 economic factors. Table 5 below summarizes the average MAE, RMSE, ND, and NRMSE values of the algorithms with their standard deviation, on the test set. First, we split the test size into batches driven by the Forecast Window parameter. Then we calculate the mean value over the non-overlapping batches and report it. We can observe that Lasso and Ridge regression models achieve the best results among all methods on spatial information. Extra trees performed the best from the tree-based models. The deep learning models show considerable improvement for multi-step predictions when we perform univariate analysis on the exchange rate target variable.

Figure 5 depicts the one-step, while figures 6 and 7 portray the multi-step predictions. We can clearly observe that the Lasso and Ridge model predictions are very close to the actual values. Additionally, for the multi-step predictions, the deep learning networks follow the trend satisfactorily with a minimal error value.

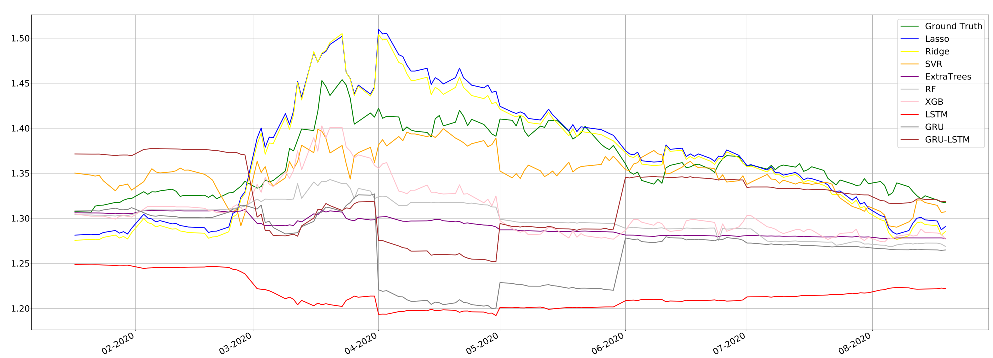


Figure 5: One-Step Model Predictions for 150 data points, Input: Features Only

4.2 Macroeconomic Features and Lag Values

This section covers the metrics when temporal information is incorporated to the existing 13 features of the dataset. Table 6 summarizes the performance of models when trained on spatial data and lag values of the exchange rate. We observe that while the performance of all the models improves with the inclusion of

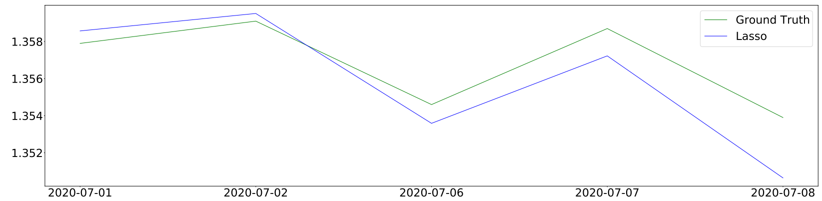
Table 5: Performance comparison of models trained on spatial data

Model	Forecast Window	MAE	ND	RMSE	NRMSE
Lasso	**1**	**0.02005 ± 0.01612**	**0.01535 ± 0.01201**	**0.02005 ± 0.01612**	**0.01535 ± 0.01201**
Ridge	**1**	**0.02373 ± 0.01586**	**0.01818 ± 0.01184**	**0.02373 ± 0.01586**	**0.01818 ± 0.01184**
SVR	1	0.07215 ± 0.04613	0.0569 ± 0.03746	0.07215 ± 0.04613	0.0569 ± 0.03746
ExtraTrees	1	0.03665 ± 0.0343	0.02733 ± 0.02428	0.03665 ± 0.0343	0.02733 ± 0.02428
RF	1	0.04426 ± 0.02948	0.03355 ± 0.02146	0.04426 ± 0.02948	0.03355 ± 0.02146
XGB	1	0.05643 ± 0.03883	0.04343 ± 0.03015	0.05643 ± 0.03883	0.04343 ± 0.03015
LSTM	1	0.07213 ± 0.06515	0.05374 ± 0.04642	0.07213 ± 0.06515	0.05374 ± 0.04642
GRU	1	0.05872 ± 0.04764	0.04409 ± 0.03372	0.05872 ± 0.04764	0.04409 ± 0.03372
GRU-LSTM	1	0.04154 ± 0.03584	0.03162 ± 0.02589	0.04154 ± 0.03584	0.03162 ± 0.02589
Lasso	5	0.02009 ± 0.01482	0.01539 ± 0.01103	0.02098 ± 0.01493	0.01606 ± 0.01107
Ridge	5	0.02386 ± 0.01468	0.01828 ± 0.01096	0.02465 ± 0.0147	0.01888 ± 0.01095
SVR	5	0.07237 ± 0.04548	0.05708 ± 0.03699	0.07315 ± 0.04493	0.05765 ± 0.03658
ExtraTrees	5	0.03662 ± 0.03384	0.02731 ± 0.02394	0.03715 ± 0.03366	0.02774 ± 0.0238
RF	5	0.04438 ± 0.02883	0.03365 ± 0.02099	0.04491 ± 0.02867	0.03407 ± 0.02086
XGB	5	0.05672 ± 0.03827	0.04366 ± 0.02976	0.05731 ± 0.03815	0.04413 ± 0.02967
LSTM	5	0.00598 ± 0.00399	0.00453 ± 0.00302	0.00598 ± 0.00399	0.00453 ± 0.00302
GRU	5	0.00666 ± 0.00356	0.00504 ± 0.00268	0.00666 ± 0.00356	0.00504 ± 0.00269
GRU-LSTM	**5**	**0.005 ± 0.00316**	**0.00379 ± 0.00239**	**0.005 ± 0.00316**	**0.00379 ± 0.00239**
Lasso	10	0.02003 ± 0.01403	0.01534 ± 0.0104	0.02147 ± 0.01411	0.01644 ± 0.01042
Ridge	10	0.0238 ± 0.01372	0.01823 ± 0.01021	0.02505 ± 0.01379	0.01919 ± 0.01024
SVR	10	0.07235 ± 0.04484	0.05706 ± 0.03652	0.0738 ± 0.04378	0.05813 ± 0.03571
ExtraTrees	10	0.03662 ± 0.03341	0.02731 ± 0.02362	0.03762 ± 0.03313	0.02812 ± 0.02339
RF	10	0.04438 ± 0.02814	0.03365 ± 0.02045	0.04548 ± 0.02775	0.03453 ± 0.02014
XGB	10	0.05671 ± 0.03739	0.04366 ± 0.02909	0.05803 ± 0.03701	0.0447 ± 0.0288
LSTM	10	0.00798 ± 0.00506	0.00604 ± 0.00384	0.00798 ± 0.00506	0.00604 ± 0.00383
GRU	10	0.01245 ± 0.00587	0.00945 ± 0.00447	0.01245 ± 0.00587	0.00943 ± 0.00446
GRU-LSTM	**10**	**0.00722 ± 0.00435**	**0.00547 ± 0.00329**	**0.00722 ± 0.00435**	**0.00547 ± 0.00328**

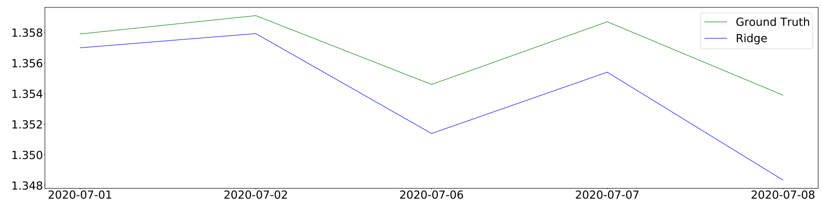
temporal information, the most significant improvement is witnessed for the LSTM architecture due to its ability to preserve long sequential data of the lag inputs.

4.3 Lag Values Only Models

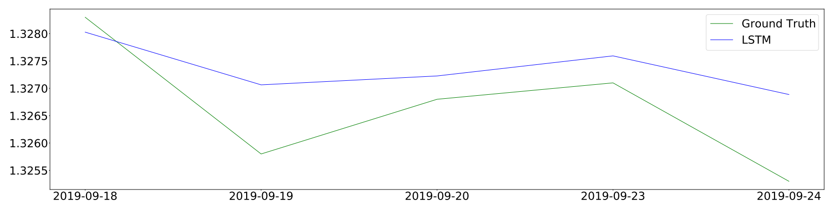
For analyzing the influence of temporal information, Table 7 summarizes the evaluation metrics of the algorithms trained on the lag values only. We can observe a boost in the performance over one-step predictions. However, there is not a significant enhancement across multi-step predictions. Figures 8, 9, and 10 depict the one-step and multi-step model predictions, respectively. The plots illustrate that the one-step predictions



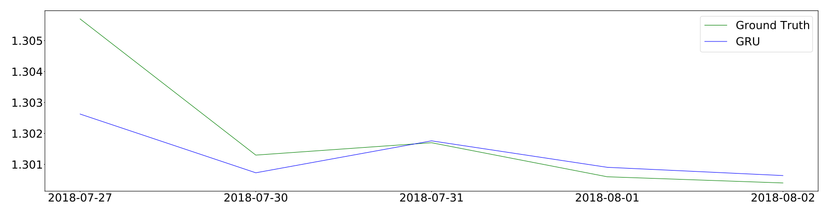
(a) Lasso



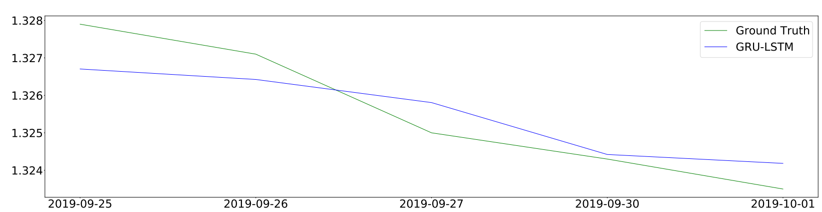
(b) Ridge



(c) LSTM

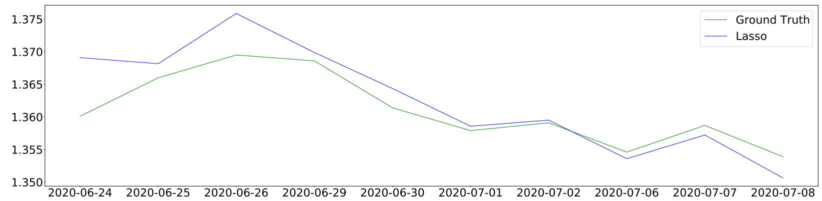


(d) GRU

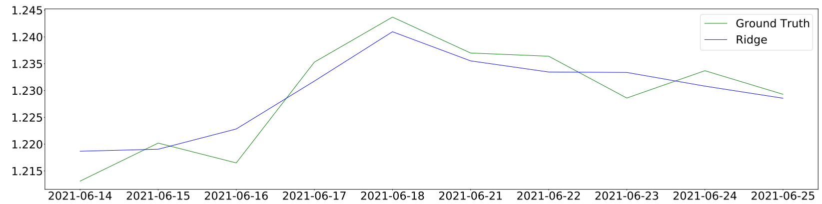


(e) GRU-LSTM

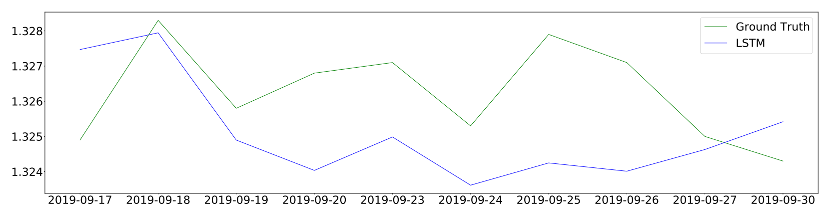
Figure 6: 5-Step Model Predictions, Input: Features Only



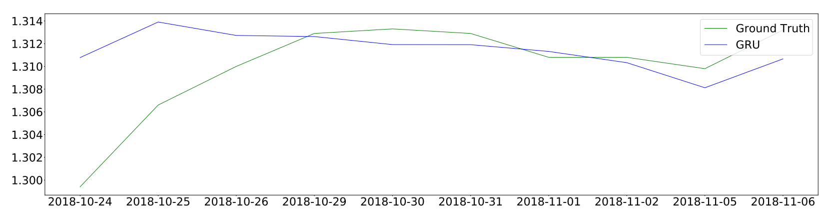
(a) Lasso



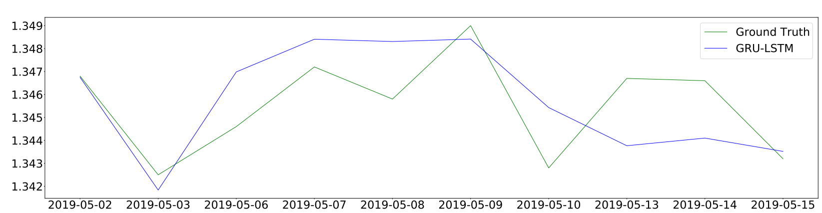
(b) Ridge



(c) LSTM



(d) GRU



(e) GRU-LSTM

Figure 7: 10-Step Model Predictions, Input: Features Only

Table 6: Performance comparison of RNN models trained on spatial and temporal data

Model	Forecast Window	MAE	ND	RMSE	NRMSE
LSTM	**1**	**0.00761 ± 0.00704**	**0.0058 ± 0.00515**	**0.00761 ± 0.00704**	**0.0058 ± 0.00515**
GRU	1	0.00846 ± 0.008	0.00646 ± 0.00582	0.00846 ± 0.008	0.00646 ± 0.00582
GRU-LSTM	1	0.00808 ± 0.00758	0.00615 ± 0.00553	0.00808 ± 0.00758	0.00615 ± 0.00553

are very well-matched with the true values. Furthermore, the multi-step values follow the general trend of the actual exchange rate movement.

Table 7: Performance comparison of RNN models trained on temporal data

Model	Forecast Window	MAE	ND	RMSE	NRMSE
LSTM	**1**	**0.00639 ± 0.00628**	**0.00485 ± 0.00452**	**0.00639 ± 0.00628**	**0.00485 ± 0.00452**
GRU	1	0.00772 ± 0.00805	0.00581 ± 0.00573	0.00772 ± 0.00805	0.00581 ± 0.00573
GRU-LSTM	1	0.00654 ± 0.00606	0.00499 ± 0.00444	0.00654 ± 0.00606	0.00499 ± 0.00444
LSTM	5	0.01079 ± 0.00786	0.00829 ± 0.00577	0.01079 ± 0.00786	0.00829 ± 0.00578
GRU	**5**	**0.00792 ± 0.00681**	**0.00602 ± 0.00486**	**0.00792 ± 0.00681**	**0.00603 ± 0.00487**
GRU-LSTM	5	0.0089 ± 0.00738	0.00676 ± 0.00527	0.0089 ± 0.00738	0.00677 ± 0.00528
LSTM	10	0.01293 ± 0.01101	0.00979 ± 0.00778	0.01293 ± 0.01101	0.00981 ± 0.00782
GRU	**10**	**0.01036 ± 0.00818**	**0.0079 ± 0.00586**	**0.01036 ± 0.00818**	**0.00791 ± 0.0059**
GRU-LSTM	10	0.0108 ± 0.00872	0.0082 ± 0.0062	0.0108 ± 0.00872	0.00821 ± 0.00625

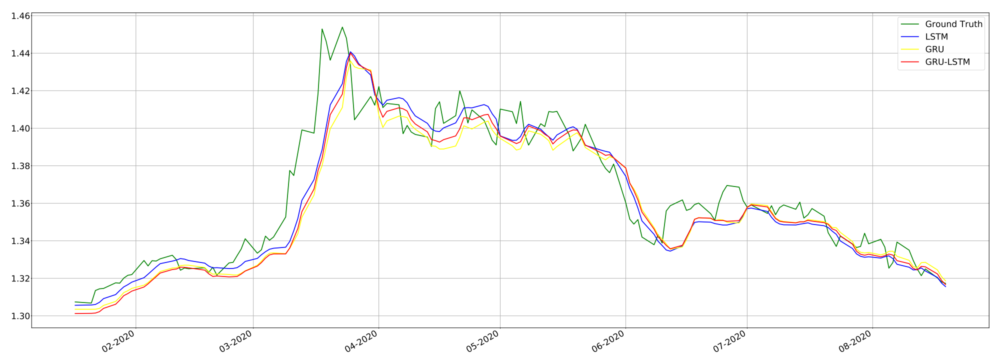
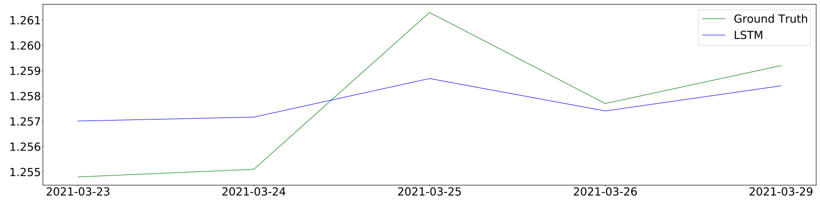
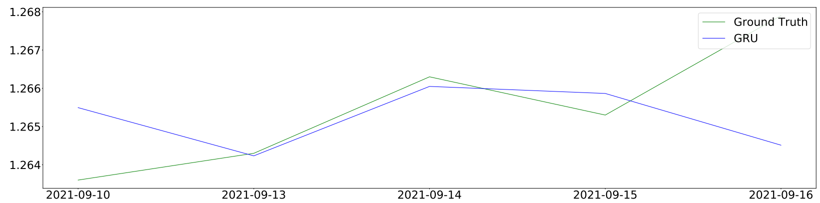


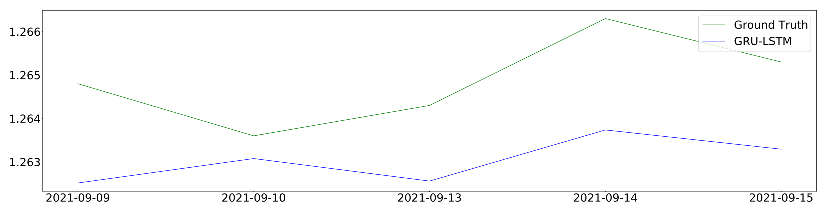
Figure 8: One-Step Lag Model Predictions for 150 data points, Input: Lag Values Only



(a) LSTM



(b) GRU



(c) GRU-LSTM

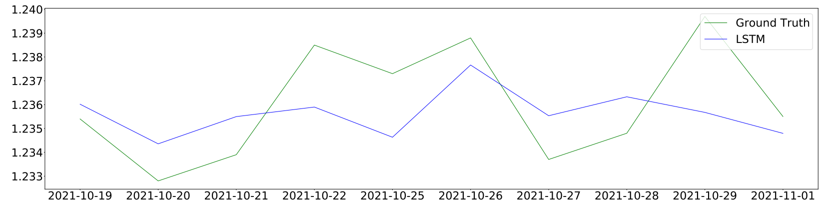
Figure 9: 5-Step Lag Model Predictions, Input: Lag Values Only

4.4 Feature Importance Estimation

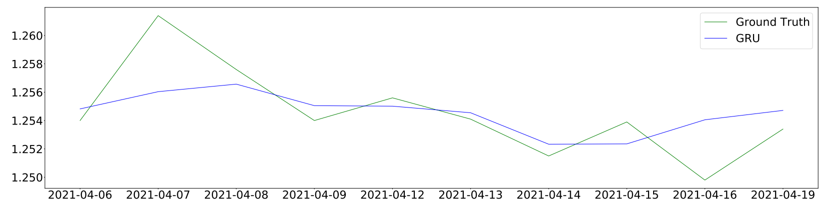
In this section, we analyze the significance of the economic indicators in driving the foreign exchange rate value. We first look at the model coefficients for the best performing traditional models, followed by comparing the feature importance of tree-based models. Lastly, we understand the impact of the features in the case of deep-learning model predictions.

Figure 11 represent the feature importance of the Lasso and Ridge Regression models, respectively. We note that Purchasing Power, Industrial Production, Unemployment, Oil, and Money supply stocks contribute the highest significance to the model predictions.

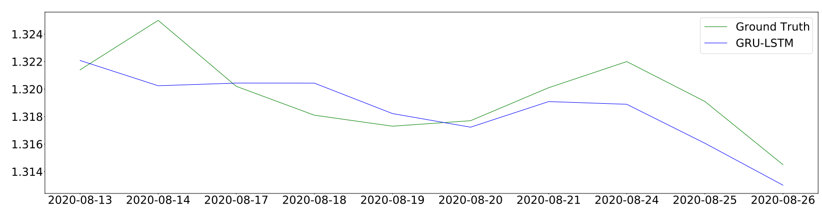
Figures 12, 13, and 14 illustrate the feature importance of the RF, ET, and XGBoost models, respectively. The plots use the in-built feature importance property of the models and the Tree SHAP technique applied to each model. The in-built property calculates the significance based on the training data, while Tree SHAP utilizes the testing data for computation. Using both techniques provides a comprehensive view of how each attribute contributes to the predictions. Considering both the feature importance plots, we perceive that although the magnitude and order differ, Money supply stocks, Oil, Commodity Price Index, and S&P500



(a) LSTM

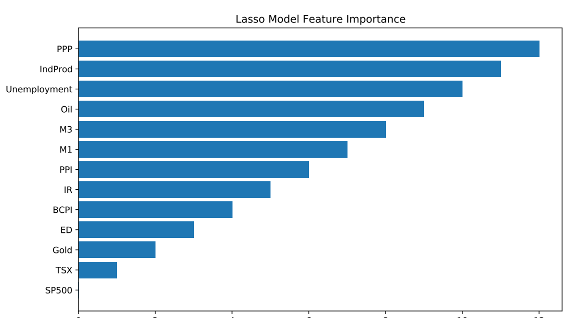


(b) GRU

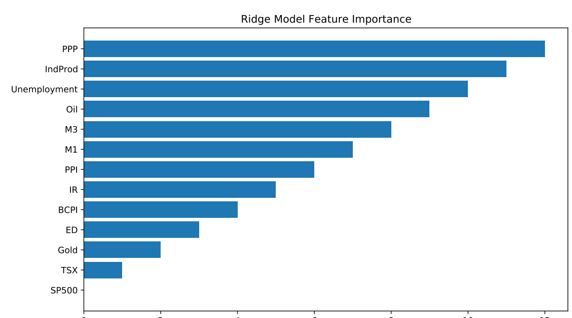


(c) GRU-LSTM

Figure 10: 10-Step Lag Model Predictions, Input: Lag Values Only



(a) Lasso

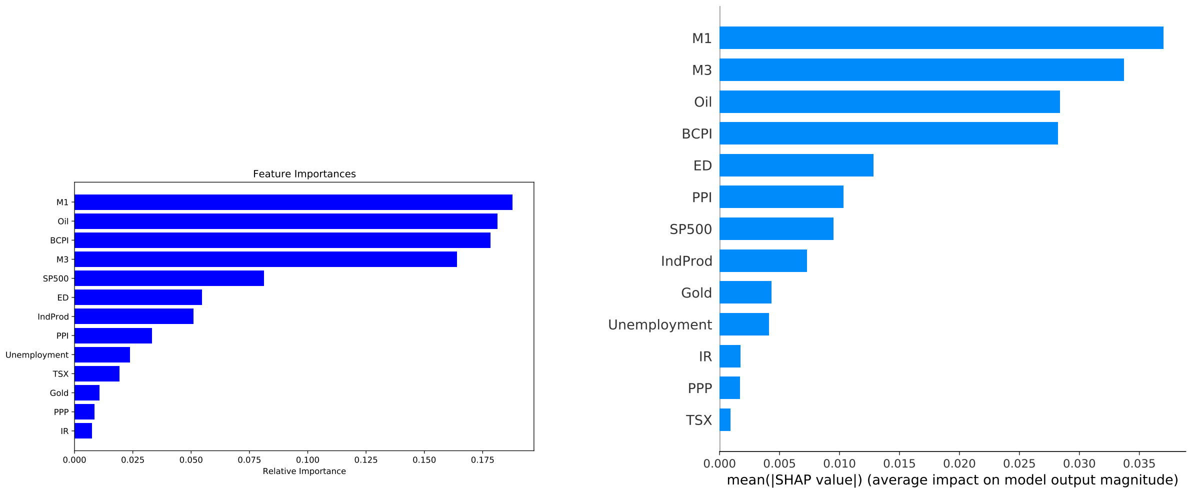


(b) Ridge

Figure 11: Lasso and Ridge Model Coefficients - Features Only

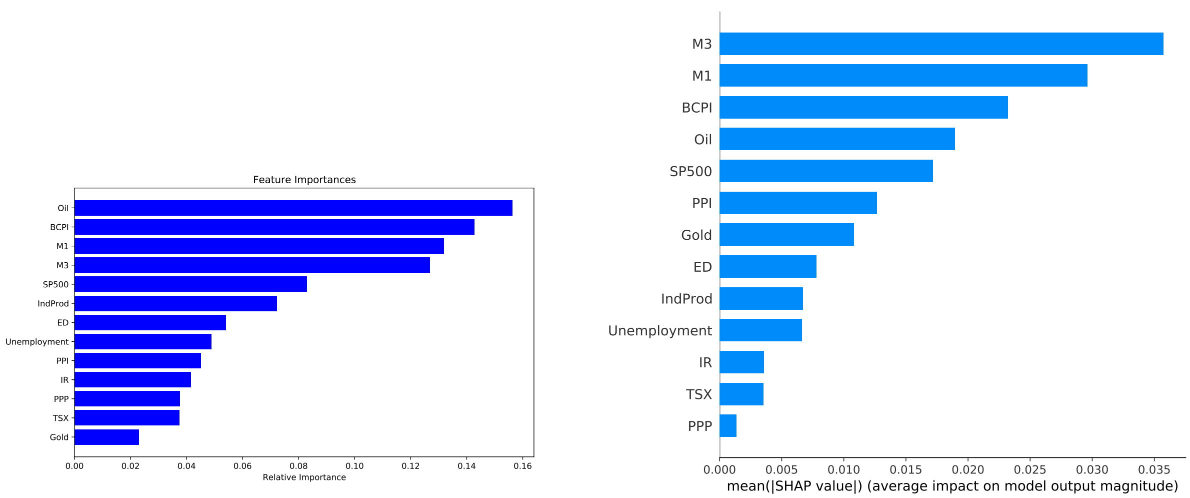
Index are the top significant fundamentals.

Figure 15 represents the feature importance of the RNNs trained on spatial information only. We can



(a) RF Feature Importance (b) RF Tree SHAP

Figure 12: Random Forest Feature Importance - Features Only

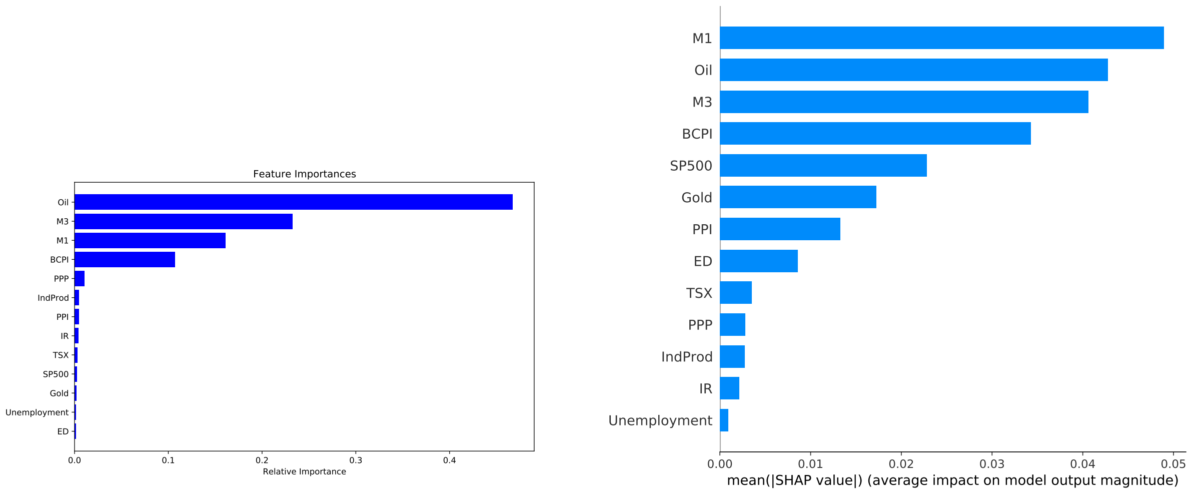


(a) ET Feature Importance (b) ET Tree SHAP

Figure 13: Extra Trees Feature Importance - Features Only

observe that all the deep learning networks consider M1, M3, BCPI & Oil, have a predominant impact on the predictions with respect to the other factors.

Figure 16 illustrates the impact of temporal information in the feature importance charts, when introduced along with spatial data. We see the influence of lag values here, with T - 4 day's lag value being the most significant than the others. Finally, Figure 17 portrays the importance of the RNNs trained on temporal information only. As anticipated, the T - 1 day's lag value turned out to be the most significant, following T - 2, T -3, and so on.



(a) XGB Feature Importance

(b) XGB Tree SHAP

Figure 14: XGBoost Feature Importance - Features Only

5 Conclusions

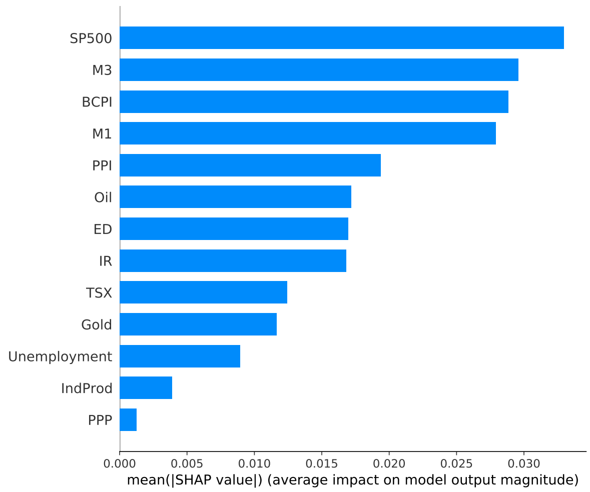
Although predicting foreign exchange rate movement with accuracy has been a concern and a difficult task for legislators, this remains the most crucial problem in the literature. Determining macroeconomic variables that are responsible for the change in currency direction is also an important issue. This research studied the causal connection between several macroeconomic variables and the USD-CAD exchange rate. We have compared the forecasting performance of the traditional linear, tree-based, and RNN models, examining the impact of introducing temporal information to the dataset. From the results of one-day predictions, the Linear Regression models seem to perform better than other models. Conversely, the RNN models have an edge when temporal information is incorporated for multi-step predictions. The Money stocks, Oil, S&P 500 Index, Purchasing Power, and Unemployment were the most significant factors generally.

Scholars and practitioners place great importance on having access to high-frequency data for current developments in modeling and forecasting foreign exchange rates utilizing intraday information. The financial community had a boost lately with the introduction of deep learning implementations for prediction research in the economic domain, and many recent publications have appeared accordingly. In our survey, we wanted to review existing studies to provide a snapshot of the current research status of conventional and deep learning implementations for financial time series forecasting.

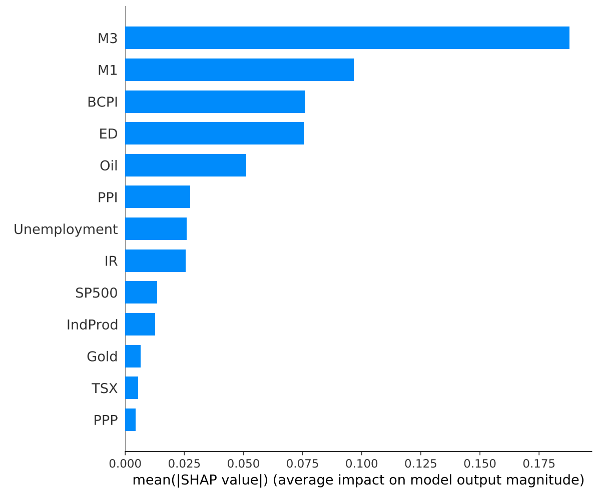
5.1 Challenges and Study Limitations

We summarize the challenges faced when building our forecasting models and the limitations present in our work as follows:

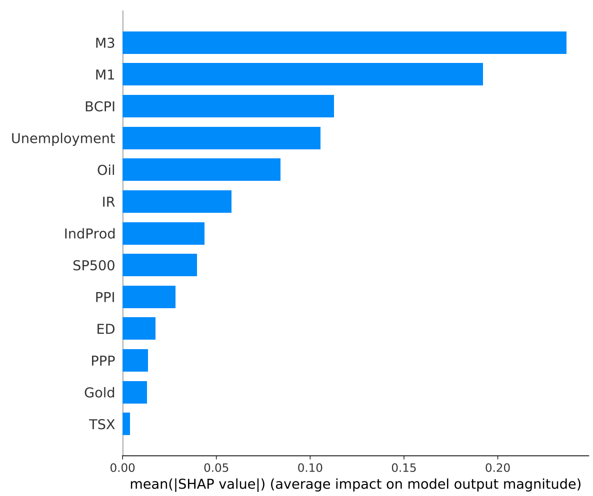
- The time series for exchange rates belong to non-stationary and non-linear data, making the task of



(a) LSTM Deep SHAP



(b) GRU Deep SHAP

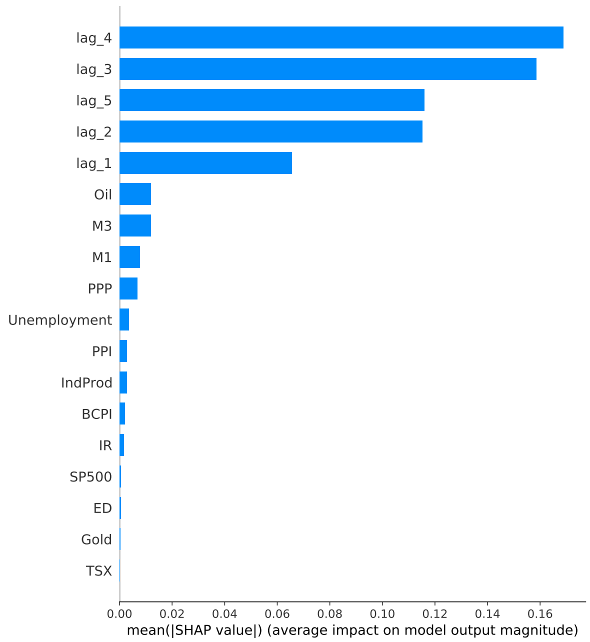


(c) GRU-LSTM Deep SHAP

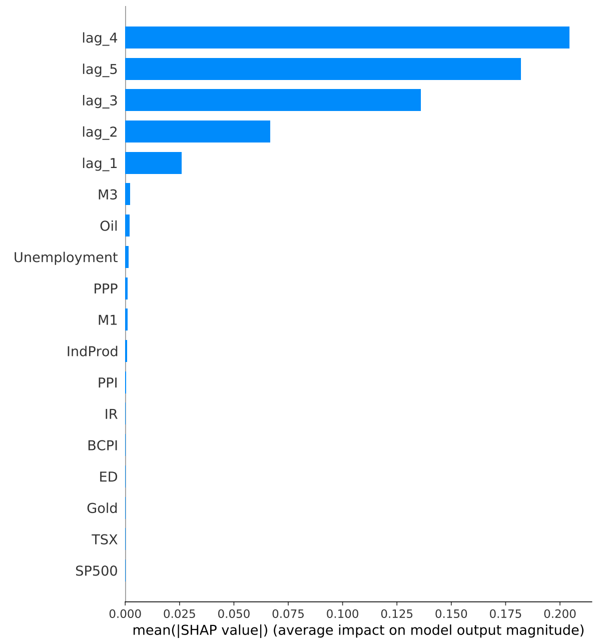
Figure 15: RNN Feature Importance - Features Only

forecasting extremely challenging.

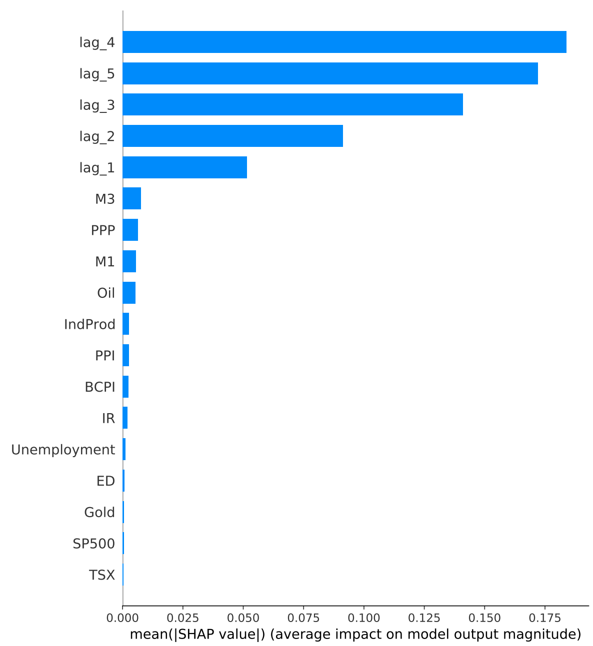
- The data sets we incorporated in this study had different frequencies, making it tedious to perform the data integration.
- Due to the massive scale of our input data combined from multiple sources, it became difficult to perform extensive hyperparameter tuning and experiment with more complex models.



(a) LSTM Deep SHAP

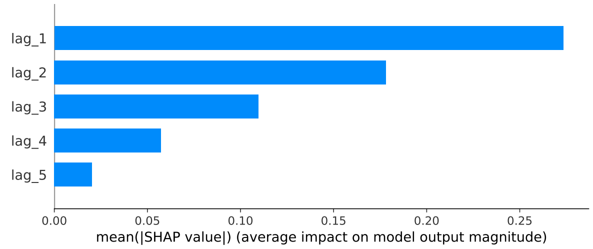


(b) GRU Deep SHAP

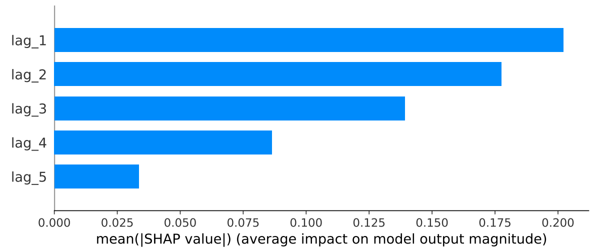


(c) GRU-LSTM Deep SHAP

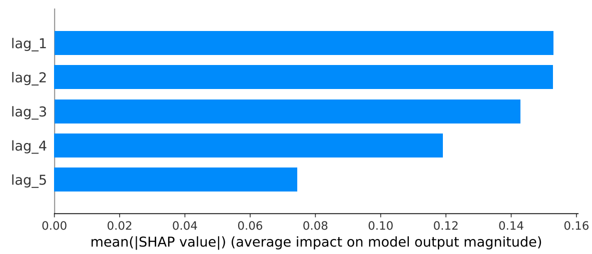
Figure 16: RNN Feature Importance - Features + Lag Values



(a) LSTM Deep SHAP



(b) GRU Deep SHAP



(c) GRU-LSTM Deep SHAP

Figure 17: RNN Feature Importance - Lag Values Only

5.2 Future Research Directions

Following future research directions can be considered as an extension of the research conducted in this project:

- We suggest that researchers may obtain more accurate results and make more precise predictions by using more macroeconomic variables.
- This study can be extended to a wider period covering the major crises experienced in the 1990s to see the prediction capabilities of different methods.
- By extending this study with more variables covering a larger time interval, a potential global crisis can be predicted before its occurrence. An augmented out-of-sample technique can be employed to improve the prediction performance for the test batches that are farthest from the training set.

References

[1] Srihari Athiyarath, Mousumi Paul, and Srivatsa Krishnaswamy. A comparative study and analysis of time series forecasting techniques. *SN Computer Science*, 1(3):1–7, 2020.

[2] Alev Dilek Aydin and Seyma Caliskan Cavdar. Comparison of prediction performances of artificial neural network (ann) and vector autoregressive (var) models by using the macroeconomic variables of gold prices, borsa istanbul (bist) 100 index and us dollar-turkish lira (usd/try) exchange rates. *Procedia Economics and Finance*, 30:3–14, 2015.

[3] Joscha Beckmann and Rainer Schüssler. Forecasting exchange rates under parameter and model uncertainty. *Journal of International Money and Finance*, 60:267–288, 2016.

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

[5] Gabriel Bruneau and Kevin Moran. Exchange rate fluctuations and labour market adjustments in canadian manufacturing industries. *Canadian Journal of Economics/Revue canadienne d'économique*, 50(1):72–93, 2017.

[6] Joseph P Byrne, Dimitris Korobilis, and Pinho J Ribeiro. Exchange rate predictability in a changing world. *Journal of International Money and Finance*, 62:1–24, 2016.

[7] Guillermo A Calvo and Carlos Alfredo Rodriguez. A model of exchange rate determination under currency substitution and rational expectations. *Journal of Political Economy*, 85(3):617–625, 1977.

[8] Shutao Cao, Wei Dong, and Ben Tomlin. The sensitivity of producer prices to exchange rates: insights from micro data. Technical report, Bank of Canada Working Paper, 2012.

[9] Yin-Wong Cheung, Menzie D Chinn, and Antonio Garcia Pascual. Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24(7):1150–1175, 2005.

[10] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[11] John H Cochrane. Time series for macroeconomics and finance. *Manuscript, University of Chicago*, pages 1–136, 2005.

[12] Christian D Dick, Ronald MacDonald, and Lukas Menkhoff. Exchange rate forecasts and expected fundamentals. *Journal of International Money and Finance*, 53:235–256, 2015.

[13] Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017.

[14] Roberto Frenkel and Jaime Ros. Unemployment and the real exchange rate in latin america. *World Development*, 34(4):631–646, 2006.

[15] Jaqueson K Galimberti and Marcelo L Moura. Taylor rules and exchange rate predictability in emerging economies. *Journal of International Money and Finance*, 32:1008–1031, 2013.

[16] Mohammadreza Ghadimpour and Seyed babak Ebrahimi. Forecasting financial time series using deep learning networks: Evidence from long-short term memory and gated recurrent unit. *Iranian Journal of Finance*, pages –, 2022. ISSN 2676-6337. doi: 10.30699/ijf.2022.313164.1286. URL http://www.ijfifsa.ir/article_150173.html.

[17] Yixin Guo. Stock price prediction using machine learning, 2022.

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

[19] Md Saiful Islam and Emam Hossain. Foreign exchange currency rate prediction using a gru-lstm hybrid network. *Soft Computing Letters*, page 100009, 2020.

[20] SFN Islam, A Sholahuddin, and AS Abdullah. Extreme gradient boosting (xgboost) method in making forecasting application and analysis of usd exchange rates against rupiah. In *Journal of Physics: Conference Series*, volume 1722, page 012016. IOP Publishing, 2021.

[21] Aikaterini Karanikola, Charalampos M Liapis, and Sotiris Kotsiantis. A comparison of contemporary methods on univariate time series forecasting. In *Advances in Machine Learning/Deep Learning-based Technologies*, pages 143–168. Springer, 2022.

[22] Ritvik Khandelwal, Prasham Marfatia, Shubham Shah, Varun Joshi, Pratik Kamath, and Kshitij Chavan. Financial data time series forecasting using neural networks and a comparative study. In *2022 International Conference for Advancement in Technology (ICONAT)*, pages 1–6, 2022. doi: 10.1109/ICONAT53423.2022.9725845.

[23] Been Kim, Rajiv Khanna, and Oluwasanmi O Koyejo. Examples are not enough, learn to criticize! criticism for interpretability. *Advances in neural information processing systems*, 29, 2016.

[24] Roy Kouwenberg, Agnieszka Markiewicz, Ralph Verhoeks, and Remco CJ Zwinkels. Model uncertainty and exchange rate forecasting. *Journal of Financial and Quantitative Analysis*, 52(1):341–363, 2017.

[25] Bjoern Krollner, Bruce J Vanstone, Gavin R Finnie, et al. Financial time series forecasting with machine learning techniques: a survey. In *ESANN*, 2010.

[26] Ruofan Liao, Petchaluck Boonyakunakorn, Napat Harnpornchai, and Songsak Sriboonchitta. Forecasting the exchange rate for usd to rmb using rnn and svm. In *Journal of Physics: Conference Series*, volume 1616, page 012050. IOP Publishing, 2020.

[27] Shixia Liu, Xiting Wang, Mengchen Liu, and Jun Zhu. Towards better analysis of machine learning models: A visual analytics perspective. *Visual Informatics*, 1(1):48–56, 2017.

[28] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

[29] Paravee Maneejuk and Wilawan Srichaikul. Forecasting foreign exchange markets: further evidence using machine learning models. *Soft Computing*, 25(12):7887–7898, 2021.

[30] Tanya Molodtsova and David H Papell. Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of International Economics*, 77(2):167–180, 2009.

[31] Manaswinee Madhumita Panda, Surya Narayan Panda, and Prasant Kumar Pattnaik. Multi currency exchange rate prediction using convolutional neural network. *Materials Today: Proceedings*, 2021.

[32] Sangjin Park and Jae-Suk Yang. Interpretable deep learning lstm model for intelligent economic decision-making. *Knowledge-Based Systems*, 248:108907, 2022.

[33] Juan Carlos Reboredo, Miguel A Rivera-Castro, and Gilney F Zebende. Oil and us dollar exchange rate dependence: A detrended cross-correlation approach. *Energy Economics*, 42:132–139, 2014.

[34] Marko Robnik-Šikonja and Marko Bohanec. Perturbation-based explanations of prediction models. In *Human and machine learning*, pages 159–175. Springer, 2018.

[35] Barbara Rossi. Are exchange rates really random walks? some evidence robust to parameter instability. *Macroeconomic Dynamics*, 10(1):20–38, 2006.

[36] Pradeepta Kumar Sarangi, Muskaan Chawla, Pinaki Ghosh, Sunny Singh, and PK Singh. Forex trend analysis using machine learning techniques: Inr vs usd currency exchange rate using ann-ga hybrid approach. *Materials Today: Proceedings*, 49:3170–3176, 2022.

[37] Omer Berat Sezer, Mehmet Ugur Gudelek, and Ahmet Murat Ozbayoglu. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90: 106181, 2020.

[38] L Shapley. 7. a value for n-person games. contributions to the theory of games ii (1953) 307-317. In *Classics in Game Theory*, pages 69–79. Princeton University Press, 2020.

[39] Hua Shen and Xun Liang. A time series forecasting model based on deep learning integrated algorithm with stacked autoencoders and svr for fx prediction. In *International Conference on Artificial Neural Networks*, pages 326–335. Springer, 2016.

[40] Alan C Stockman. A theory of exchange rate determination. *Journal of political Economy*, 88(4): 673–698, 1980.

[41] Yusui Tang, Xiao Xiao, MIM Wahab, and Feng Ma. The role of oil futures intraday information on predicting us stock market volatility. *Journal of Management Science and Engineering*, 6(1):64–74, 2021.

[42] Jonathan H Wright. Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics*, 146(2):329–341, 2008.

[43] Mo Yang and Jing Wang. Adaptability of financial time series prediction based on bilstm. *Procedia Computer Science*, 199:18–25, 2022. ISSN 1877-0509. doi: https://doi.org/10.1016/j.procs.2022.01.003. URL https://www.sciencedirect.com/science/article/pii/S1877050922000035. The 8th International Conference on Information Technology and Quantitative Management (ITQM 2020 & 2021): Developing Global Digital Economy after COVID-19.