A Guidance for The Central Bank's Monetary Policy: Using Machine Learning to Predict Inflation

Ente Kang

Introduction

With the outbreak of the COVID-19 pandemic, economies all around the world have all undertaken some form of monetary and fiscal policies to ensure the core function of their financial functions and economies remains intact [1]. The Bank of Canada (BoC) has been forecasting inflation using the ToTEM-III model and has been successful in maintaining inflation in 2020 at the lower bound of the target range, at around 2% [4]. The BoC’s Monetary policy report of April 2021 anticipates heightened, but transitory inflation numbers in the second quarter of 2021

Our motivation arises from central banks’ reluctance to incorporate machine learning models in their research process, whereby, the Federal Reserve has yet to centralize the use of machine learning (ML) [11]. Since central banks need to convince the readers with their reports, the lack of interpretability of machine learning models, also considered black box models, makes it difficult to do so.

This research project will try to investigate how accurately a simple machine learning model performs against more traditional forecasting techniques. It is hypothesized that machine learning models can yield accurate results or even outperform traditional forecasting methods, because the former are able to discover nuances and seemingly hidden relationships between variables that traditional methods are not capable of.

The contribution of this report stems from analyzing the benefits and downsides of incorporating ML models into central banks' analysis to provide some insights on whether it is appropriate for central banks to incorporate machine learning models in their research. Moreover, the report aims to determine whether consumer sentiment regarding inflation can be used to derive appropriate predictors to measure it since stagflation concerns have risen amidst supply bottlenecks and rising oil prices, whereas 2021 has seen somewhat exaggerated rate with headline inflation reaching 4.7% on a year-over-year basis in October.

The traditional forecasting model used in this report is vector autoregression of order 1(VAR). The ML models used are supervised learning techniques, since the goal of the ML models are to predict inflation, which is a quantitative response variable. The non-linearity characterized in financial data, render us to use ensemble methods and KNN.

The set of literature reviewed comprises of three papers. The first one is from the Bank of Canada, by Greg Tkacz and Carolyn Wilikins, called Linear and Threshold Forecasts of Output and Inflation with stock and housing prices. We agree with the authors that financial market conditions can help predict inflation but are also skeptical about whether linear methods are appropriate [13]. Our report is strictly interested in predicting the inflation using supervised machine learning techniques rather than make statistical inferences on which variables can help explain inflation.

A publication by Araujo and Gaglianone, compares the performance of traditional forecasting methods to ML models such as ridge regression and random forest using macroeconomic and financial variables [15]. Araujo and Gaglianone utilized an automated algorithm for variable selection, while we will be utilizing all the variables in our data, as well as a version guided by textual data extracted from Twitter. A recent paper published by the Bank of Italy, authored by Angelico, Marcucci, Miccoli and Quarta, focused on adopting ML methods to extract twitter data to analyse consumer sentiments on inflation. The authors created an inflation expectation index, akin to our approach of harnessing textual data to determine predictors. However, time and data extraction limitations restrict us from building our own index. [12]

Data

The data used in this paper is from Jan. 03, 2017, to Nov. 5, 2021, and was constructed from datasets found on Statistics Canada, The Bank of Canada, The Federal Reserve, Nasdaq, and Twitter. There were a total of 1213 observations and 39 variables.

Our goal is to predict daily inflation figures in Canada. All available data on inflation in Canada are in a monthly format, which poses an issue when using the monthly value for every day of the month. However, we determined that the daily inflation numbers in the United States can be a very good proxy for the inflation trend in Canada (Figure 1), partly due to the countries' proximities to one another and their financial similarities.

From the plot below we see that the daily trend in CPI is very similar for Canada and the United States, leading to the decision of choosing the daily inflation rate as a proxy for Canada.

Trends in the CPI - Canada vs. USA

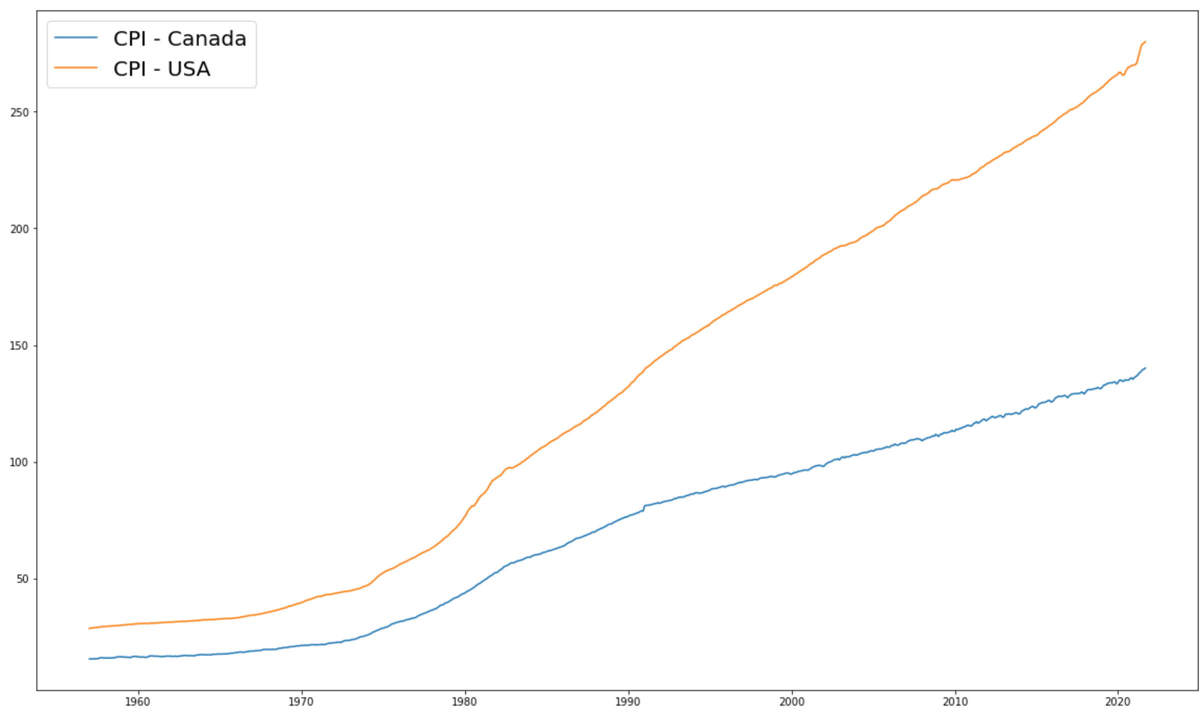

Figure 1

Numerous predictors that are included were guided by the Aggregate Demand (AD) and Aggregate Supply (AS) framework, and the variables that feed into it (Figure 2). This includes predictors such as bond rates, mortgage rates, stock prices and exchange rates.

$$AD = C + I + G + NX$$

$$AS = \pi^e + \gamma(Y - Y^P) + \rho$$

$where\ C = Consumption,\ I = Investment,\ G = Government\ Expenditures,\ NX = Net\ Exports$

$\pi^e = expected\ inflation$

$\rho = \ the\ sensitivity\ of\ inflation\ to\ the\ output\ gap$

$Y^p = \ the\ potential\ output,\ Y = output$

Inflation vs. Macroeconomic Indicators

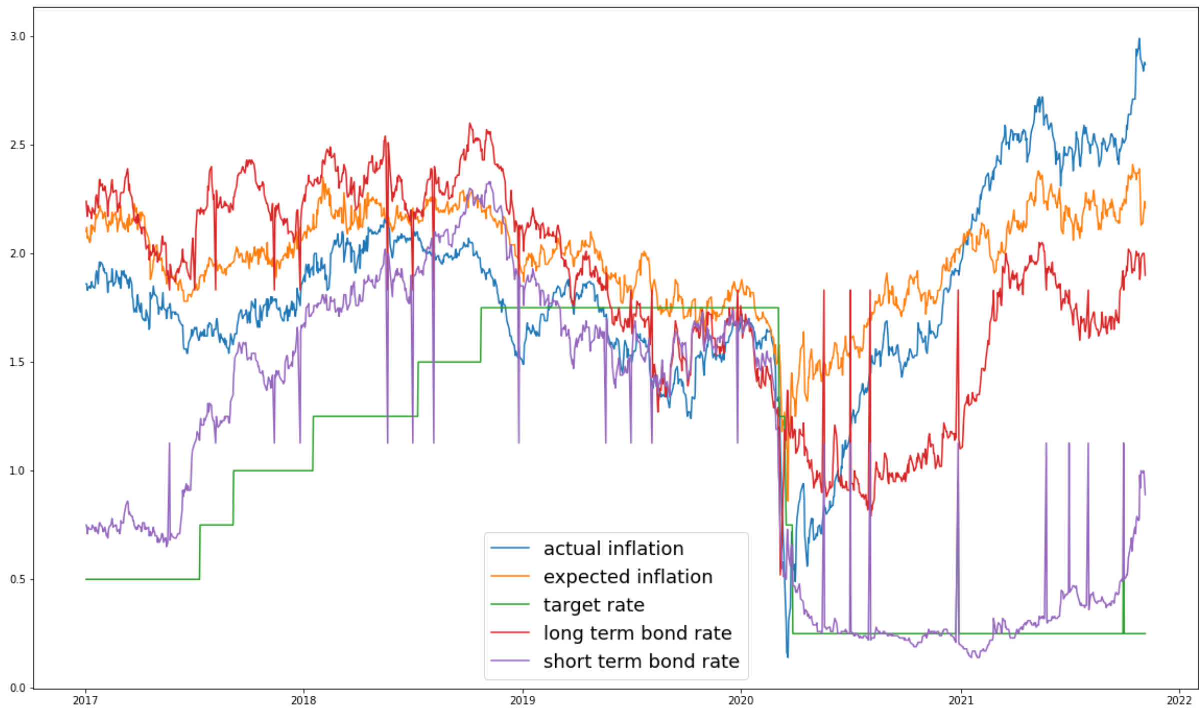

Figure 2

We have also engineered features such as a covid-19 indicator and a label for when Canada and China's relations were very tense. The reasoning is that health crises and frictions between countries can lead to changes in a country's economic outlook, which influences macroeconomic variables such as inflation.

When people behave rationally by following their expectations, those decisions manifests to real changes (Figure 3). For example, if there is an anticipated tax increase in the next quarter, households will prepare accordingly, by possibly cutting back consumption this quarter and choosing to save more during the next. When this decision-making aggregates across millions of households, it is reflected in the nation-wide consumption levels, and ultimately the inflation rate. Therefore, we have included values of households' inflation expectations as a predictor and used tweets from Twitter containing the #inflation, #outlook keywords to guide further selection of variables.

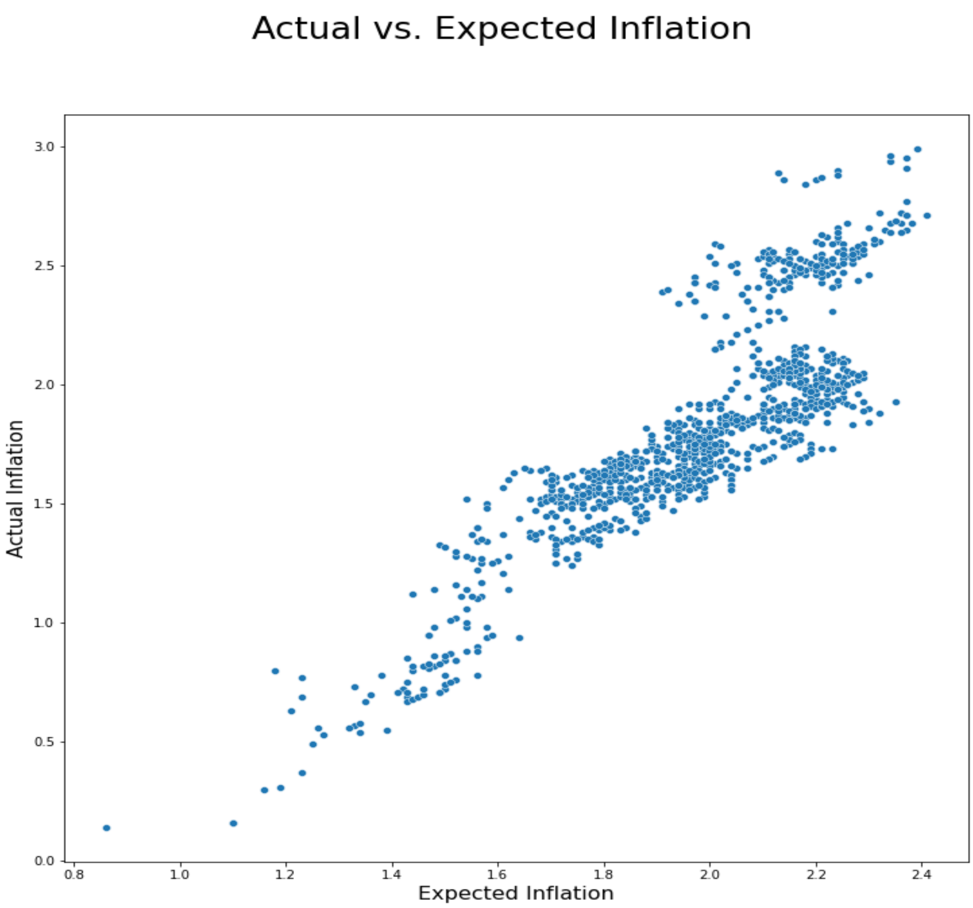

Figure 3

Methods

Forecasting was compared to ML models in this investigation. For both forecasting and ML, two different datasets were used to construct the models. A *short* version of the data excluded all the predictors that were obtained from analyzing keywords from the Twitter word cloud (Figure 4), and the *long* version included all the predictors.

The goal is to assess the difference in performance between time-series forecasting and simple Supervised ML Algorithms. However, traditional forecasting methods do not have a Mean Squared Error (MSE) metric, and can only predict future values, which makes direct comparisons with ML models difficult.

To overcome this limitation, we performed a chronological train-test split, i.e., the period from Jan. 03, 2017 to Dec. 31, 2020, served as our training data, and everything from Jan. 03, 2021 to Nov. 05, 2021, was our test data. We were then able to forecast future values, using the VAR(1) model suggested by AIC, assuring stationarity by taking the first difference, and comparing them to the true values, which leads to the creation of a Mean Squared Error metric. Several machine learning models were trained and evaluated using these exact train-test splits, helping us see if there are gains to deploying ‘blackboxes’.

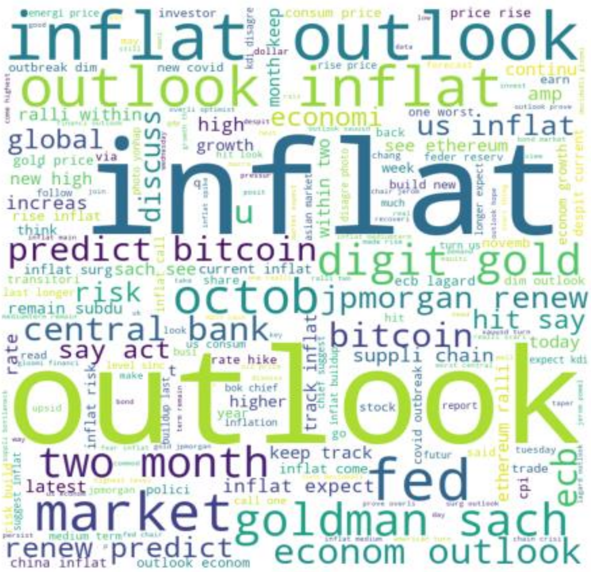

Figure 2

Lastly, a separate, random train-test split was used, with the goal of seeing which supervised machine learning model is the top-performer. The K-Nearest Neighbors Regressor, Decision Trees Regressor, and Random Forests Regressor were built to account for the non-linear structure of the movement in inflation. In all instances of our ML models, dates were removed and replaced by a covid-19 indicator and a label for times of tension between Canada, China, and the United States. The above is summarized in the following table (Table 1).

Method	Train-test split	Data
Forecasting (VAR)	Chronological	Short
	Chronological	Long
Machine Learning	Chronological	Short
	Chronological	Long
	Random	Short
	Random	Long

Table 1

Results

Our regression models - K-Nearest Neighbors (KNN), Decision Trees and Random Forests - were trained on all versions of the dataset. After running cross validation, it was determined that the easily interpretable KNN performed best. Including the predictors that were constructed based on keywords that had a high frequency in the Twitter word cloud improved the accuracy of the model in all cases. As seen below, there was a reduction in the Mean Squared Error by 54.65% (0.00094) when the train-test split was done conventionally. Even for a chronological train test split, there was a reduction in the MSE by 32.11% (0.114). In every scenario, the simple and intuitive KNN algorithm managed to outperform traditional forecasting using VAR(1) (Table 2).

Method	Number of neighbors - optimized	Train-test split	Data	Mean Squared Error (MSE)
Forecasting (VAR)		Chronological	Short	0.519
		Chronological	Long	4.231
K-Nearest Neighbors	2	Chronological	Short	0.469
Regressor	3	Chronological	Long	0.355
	3	Random	Short	0.003
	2	Random	Long	0.002

Table 2

Discussion

A ML model with a low MSE needs to achieve either both low variance and bias or have the reduction of one metric be greater than the increase of the other - known as the bias variance trade-off. In our testing, the models that incorporated sentiments always performed better, which shows the importance of using households' sentiments in guiding predictions. Directly comparing VAR vs. KNN for the same train-test splits and data used, we can conclude that there are gains to using ML. There are numerous reasons to why KNN outperformed the VAR: 1) The movement in inflation is non-linear, 2) The exact date and time is irrelevant for predictions, and 3) KNN is more capable of picking up hidden trends.

KNN is a non-parametric approach, meaning that we do not make any assumptions on the functional form on *f* in y = *f*(x) + ϵ. In contrast, a pth order VAR assumes the following functional form:

$$y_t = A_0 + \sum_{\{i=1\}}^{\{p\}} A_i y_{\{t-i\}} + \sum_{\{i=1\}}^{\{p\}} ß_i x_{\{t-i\}} + \epsilon_t, \text{ where } E(\epsilon_t) = 0, and\ E(e_t e_{\{t-i\}}) = 0$$

This is a linear structure, and the last assumption tells us that the model presumes that there is no correlation over time, which is clearly false for financial data. KNN is more flexible and works better in our situation.

Another possibility for why KNN is outperforming VAR is simply the fact that specific dates do not add much valuable information to the model. In our KNN model, all dates were removed and replaced with indicators signaling major events such as the onset of covid-19. More information is contained in macroeconomic indicators, market prices and consumer sentiments than the specific date.

Building upon what was mentioned above, there must be some trend that KNN is able to pick up, that the VAR was not able to. Perhaps when people are worried about inflation, more searches about inflation related topics take place on google, which is then feeding into changes in consumption and saving decisions, culminating in changes to inflation.

Monetary officials are in general reluctant in deploying 'blackbox' methods, but our results show that there are indeed gains to using these 'blackboxes',

Limitations and Next Steps

There is no best model, only a better one. The promising results above come with limitations that need to be addressed. Our access to Twitter data was limited to the past 7 days of tweets, and full historical access is available to qualified academia professionals with higher research degrees. ML models benefit from more historical data, and textual data is no exception. Extra care is needed when working with tweets, or textual data from any online platform to avoid incorporating biases into the model. It may very well be that individuals who are posting online have more extreme views than most of the population. The algorithm will pick up these biases, which can contaminate the results. Analogously, certain 'powerful' individuals or corporations can tilt the scales more than the common forum user, future researchers can consider assigning weights to these posts so that the model is not dependent on these powerhouses.

We saw that the macroeconomic indicators in the United States can be a very good proxy for their counterparts in Canada. The fact is, they are not ‘identical’, and results should be taken with a healthy dose of skepticism.

To further exploit the positive feedback from our models, we can include more historical tweets, and even include sentiments from other platforms such as Reddit. It will also be in our favor to carefully construct a daily inflation index for Canada.

Conclusion

Monetary officials are in general reluctant to deploy ‘blackbox’ approaches, as they need to convey results for a broader audience. Our results show that there are gains to using these ‘blackboxes’, and these methods are in fact quite interpretable. Simple supervised machine learning techniques can capture complex relationships and outperform traditional forecasting, in this case VAR, in all scenarios. Decision making is highly dependent on both expectations and sentiments toward the subject matter. Careful use of textual information can provide further prediction accuracy and we recommend that future researchers incorporate it.

References:

1. “Policy Responses to covid19.” *IMF*, www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19.
2. “About the Bank.” *Bank of Canada,* https://www.bankofcanada.ca/about/
3. “Monetary Policy.” *Bank of Canada,* https://www.bankofcanada.ca/core-functions/monetary-policy/
4. “12-month change in the Consumer Price Index (CPI) and CPI excluding energy.” *Statistics Canada,* https://www150.statcan.gc.ca/n1/daily-quotidien/211117/cg-a001-eng.htm
5. “COVID-19: Actions to Support the Economy and Financial System.” *Bank of Canada,* https://www.bankofcanada.ca/markets/market-operations-liquidity-provision/covid-19-actions-support-economy-financial-system/
6. Media Relations. “Bank of Canada maintains policy rate and forward guidance, ends quantitative easing.” *Bank of Canada*, https://www.bankofcanada.ca/2021/10/fad-press-release-2021-10-27/
7. McGugan, Ian. “Rising prices, fragile economy resurrects stagflation as market threat.” *The Globe and Mail*, 2021.
8. “Product Market: Definitions, Graphs and Data.” *Bank of Canada*, https://www.bankofcanada.ca/rates/indicators/capacity-and-inflation-pressures/product-market-definitions/
9. Gordon, Julie. “Bank of Canada says economic slack not yet absorbed, but 'getting closer'.” *Reuters*, 2021.
10. Gordon, Julie. “Bank of Canada head signals rates could rise as soon as April 2022.” *Reuters, Reuters*, 2021.
11. “Irrational Exuberance: Machine Learning at the Federal Reserve.” *HBS Digital*, https://digital.hbs.edu/platform-rctom/submission/irrational-exuberance-machine-learning-at-the-federal-reserve/
12. “Si possono misurare le aspettative di inflazione usando Twitter?.” *BANCA D’ITALIA*, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2021/2021-1318/index.html
13. Tkacz, Greg, and Wilkins, Carolyn. “Linear and Threshold Forecasts of Output and Inflation with Stock and Housing Prices.”. *Bank of Canada*, 2006.
14. Puglia, Michael, and Tucker, Adam. “Machine Learning, the Treasury Yield Curve and Recession Forecasting.” *The Federal Reserve,* May 2020.
15. Araujo, Gustavo, and Gaglianone, Wagner. “Machine learning methods for inflation forecasting in Brazil: new contenders versus classical models.” 2020.