Inflation

Using Time Series Analysis to Understand and Predict US Inflation

Tyler Beisel
MSDS
CU Boulder
Tyler.Beisel@colorado.edu

ABSTRACT

The goal of this project is to build and compare multiple time series analysis models to forecast inflation in the United States. This will use publicly available Consumer Price Index (CPI) data that is available through the Bureau of Labor Statistics (BLS) website. This work will build upon other time series analysis work in an attempt to understand what factors affect inflation as well as to build a model that can accurately predict upcoming inflation.

Time series analysis forecasting is a well-studied problem in the data science world. This project will seek to compare the results of several well-known methods and decide which method is best suited to the inflation data that is available. The methods to be considered consist of simple methods such as moving averages, regressive methods such as ARIMA and SARIMA, multivariate methods such as Vector Auto Regressors, and modern deep learning solutions using RNNs and Transformers.

Introduction

Inflation is a problem that affects all of us. It raises the cost of goods that we need on a daily basis, and it decreases the value of the money that we have saved away. It makes our salaries functionally less and it limits the purchasing power that we have. In a modern global economy, rising inflation rates can be an epidemic. Suffice it to say, we are all familiar with the pains that come with inflation.

While it is true that we are affected by inflation, very few of us actually understand what causes inflation and how we can expect inflation to change over time. We accept it as a necessary evil without seeking to understand the trends behind it and the dependent variables that cause it. A greater understanding of these things would lead to a greater preparedness of the downsides of inflation as well as a greater understanding in potential political policies that could combat the problem. This project will seek to understand the patterns behind inflation as well as explore some of the dependent variables that effect it.

Research on inflation is a daily topic for economists. Top economists spend countless hours trying to understand and predict inflation based on a variety of factors. This project will not try to usurp the work of leading economists but will instead approach the problem from a data scientist's point of view with a very limited understanding of economics. It is my hope that this perspective will help fellow people who are new to the world of economics understand some of the trends and statistical backbones of inflation and the consumer price index.

Related Work

As mentioned above, time series analysis and forecasting is a very well-studied problem. The following guides will be used as a starting point:

- https://medium.com/analytics-vidhya/time-series-forecasting-a-complete-guide-d963142da33f
- https://medium.com/analytics-vidhya/practical-time-series-from-arima-to-deep-learning-part-1-b292b07ec6c3
- https://www.analyticsvidhya.com/blog/2018/09/multivariate-time-series-guide-forecasting-modeling-python-codes/
- https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/

- https://blog.dataiku.com/deep-learning-time-series-forecasting

My work will seek to take the guides listed above and apply / alter them to the specific inflation problem that I am working on.

Additionally, much research has been done on the factors that cause inflation. My work will also make use of the below resources that have researched the dependent variables that affect inflation rates.

- https://www.rba.gov.au/education/resources/explainers/causes-of-inflation.html#:~:text=More%20jobs%20and%20higher%20wages,to%20an%20increase%20in%20inflation.
- https://hbr.org/2022/12/what-causes-inflation
- https://www.bls.gov/opub/mlr/2023/beyond-bls/what-caused-inflation-to-spike-after-2020.htm

I will use these sources to gain a thorough understanding of inflation and the factors that affect it before beginning my data science work.

Proposed Work

My work will consist of the following:

1. Research on problem specific terminology and trends
    a. I will read the reports of economists to better understand this problem space that I am entering into. This will practically serve as my domain level expertise.
2. Exploratory Data Analysis
    a. Generating all relevant plots for CPI and potential dependent variables.
    b. Checking for any missing data
    c. Looking for potential seasonality
    d. Determining stationarity
    e. Testing for white noise
3. Model Building
    a. Build and experiment with the following model types
        i. Simple – Simple Average and Simple Moving Average
        ii. Exponentially Smoothed – Variations of Holt's method
        iii. Auto Regressive – ARIMA and SARIMA
        iv. Multivariate – Vector Auto Regressors
        v. Deep Learning - RNNs and Transformers
4. Model Evaluation
    a. Evaluate and rank the models that were build and experimented with
5. Final Visualizations
    a. Generate final predictions for the top performing models.

Evaluation

There are a few popular metrics that are used for comparing forecasting models. Specifically, I will be using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). These metrics will make it easy to compare the results of the models that were built. Additionally, plots of the predictions as well as the ground truth for the training and test datasets will be generated that will provide a qualitative sense of the success of the forecasting models. Together, the quantitative and qualitative methods will help me prove which model is best suited for forecasting inflation via Consumer Price Index.

As for the project as a whole, I will consider it a success if I grow in my understanding of applying time series analysis forecasting to real world problems. Time series analysis has always fascinated me but I have never studied it in depth. This project will be valuable to me if my understanding of the subject matter increases.

Discussion

I am doing this research project during the 2023 Fall 1 Term. Therefore, all work shall be completed before

the term submission deadline of Oct 17. Some potential challenges include possible difficulties with getting the data from the BLS website, conducting all of the experiments within the allotted time, and completing the subject matter background research in a timely fashion. If there is a delay in any of the work, the number of models to be experimented with will be reduced to still complete the project in time.

Conclusion

In summary, this project seeks to understand inflation through consumer price index and a variety of other dependent variables. A variety of time series forecasting models will be built and evaluated for the task at hand. At the conclusion of the work, I will know which model forecasts the best for this problem. Future work that could come out of this could include evaluating the performance of the predictions made for the future, searching for additional dependent variables, and contributing to the vast research on the advancement of RNNs and Transformers.

References

All references have been linked in the text where the reference occurred.