
Machine Learning Engineer Nano degree Program

Federal Reserve Interest Rates

Jovani Martinez Rico - May 22, 2017



Capstone proposal

- **Domain Background**

As a current student of a Master in Economics, I know how important is the Federal Reserve for the economy of the country, any decision by the Fed directly affects mortgage, credit card, or prices rising in the product that we buy. Therefore the Federal Reserve sets interest rates to promote conditions that achieve high employment, low and stable inflation, sustainable economic growth, and moderate long-term interest rates. The Fed monitors inflation indicators to manage inflation. When indicators rise more than 3% a year, the Fed raises the federal funds rate to keep rising prices under control.

Higher interest rates mean higher borrowing costs, so consumers and businesses borrow less and spend less. Demand for goods and services drop, and inflation falls. On the other hand, falling interest rates result when the Fed lowers the federal funds rate. Borrowing becomes cheaper and people spend more. This can end a recession.

Knowing the interest rate could help companies to decide if they invest or not, help people to get a new mortgage or help people to manage better their money.

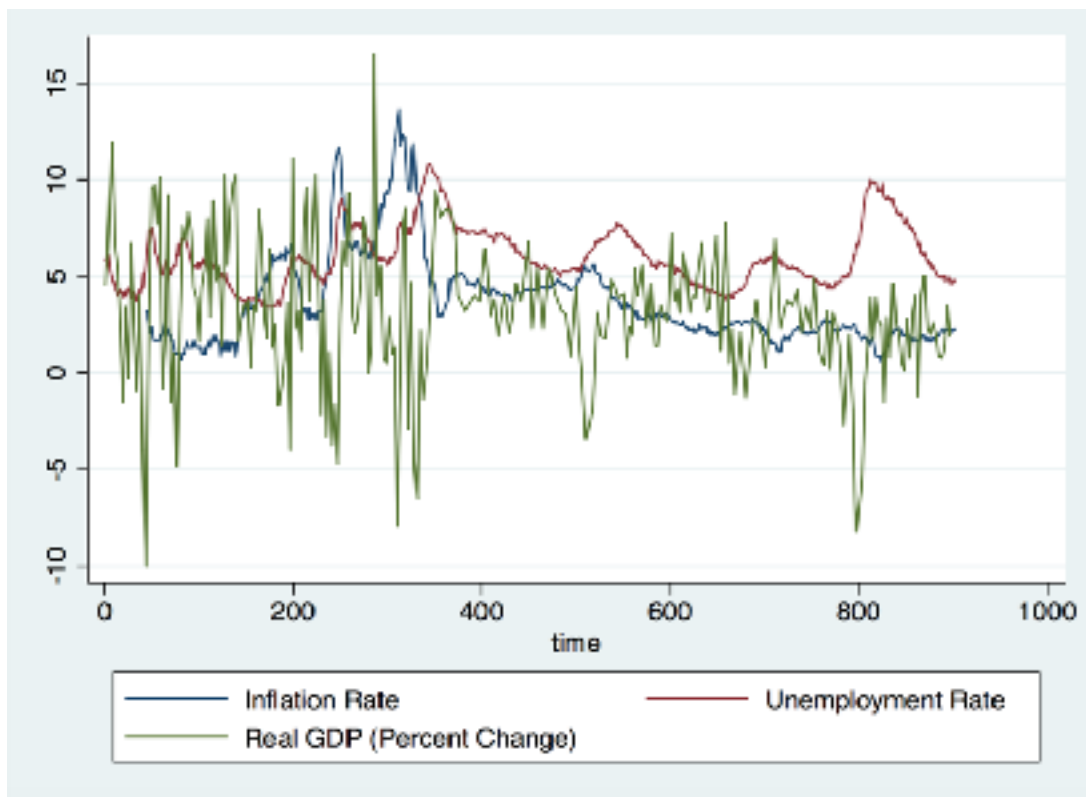
- **Problem Statement**

Looking at the country's unemployment data as well as gross domestic product and inflation rates you may try to predict the following Fed changes in the interest rate. This is a clear example of a regression problem, making a regression along of the variables we can try to predict the expecting interest rate.

Trying to predict the next interest rate can be beneficial for both a company and a person. This can help you make a better decision on their finances, they can invest their money or save it for a future.

- **Datasets and Inputs**

The data set for this project will be obtained from the Federal Reserve Bank of St. Louis' economic data portal. The consumer price index and unemployment data were provided by the US Bureau of Labor Statistics. The gross domestic product data was provided by the US Bureau of Economic Analysis. This dataset includes data on the economic conditions in the United States on a monthly basis since 1954.



- **Solution Statement**

To construct the model I will use a supervised learning algorithm, a regression is the best option to predict the next interest rate. Using the consumer price index, gross domestic product and the unemployment rate as independent variables the model will predict the interest rate.

Using different supervised algorithms will be evaluated to see which one make a better prediction, the algorithms that I will test will be KNeighborsRegressor, LinearRegression, Ridge Regression and Lasso.

- **Benchmark Model**

The Federal Reserve Bank of St. Louis' economic data portal provides information on the past interest rates that the Fed set, using that information and the variance inflation factor (VIF), we can verify the performance of the model.

- **Evaluation Metrics**

To evaluate the performance of the model, I will use the r-squared to check how well the model is getting, but also I will use the adjusted r-squared to evaluated if adding variables to the model, the accuracy of this is better with those variables, As additional explanatory variables are added to a model, r-squared will always go up, no matter how weak the relationship with Y, but the adjusted r-square introduces a “penalty” for each additional variable added. If adjusted r-square increases, then the explanatory power of the new variable is large enough to overcome the penalty, and therefore our model has improved. for each additional variable added.

- ## Project Design

1. Get the information from the Federal Reserve Bank of St. Louis' economic data portal.
2. Analyze the data if there are outliers or if the information needs any transformation (logarithm or exponential).
3. Create the model using the different algorithms (KNeighborsRegressor, LinearRegression, Ridge Regression and Lasso) and compare each one and see which one is better.
4. Check the model for multicollinearity and heteroskedasticity using F-stat and t-test.
5. Once an algorithm has been identified to has the best accuracy, use Grid Search to improved to try to make it even better.