Analysis on Inflation

Yongzhen Huang

4/18/2022

Introduction

We currently live in a special time in history. While we are still dealing with an ongoing global pandemic, the economy is also faced with high inflation. High inflation often leads to higher cost of living and decreased living standard. Understanding inflation is crucial for government and policy makers as it greatly impacts the decision making process.

This project aims to answer how inflation is tied to an economy and the economic growth, and how one may predict the future inflation given current state of the economy. In particular, the following questions will be answered

1. How does inflation affect the overall economy and economic growth (e.g. GDP)?
2. How does inflation affect the standard of living e.g. rent cost?
3. Given the current state of the economy, how may one predict the future inflation?

Methods

Data Source

The main source of data is from Organisation for Economic Co-operation and Development (OECD). Data from OECD are downloaded as annual data in CSV files and includes the following information.

Category	Details
GDP	$USD millions, per capita
Population	# million persons, growth, % working age
Employment Rate	% of working age population
Inflation	annual %
Prices	housing (rent) and share prices
Interest Rate	government short-term interest rate
Household stats	Savings and spending

Data Cleaning

Overall, the main tools used for the project are R.

R is used for the data cleaning, analytics, and constructing interactive plots

- `tidylr` for piping
- `dplyr` for table manipulations
- `data.tables` for more table manipulations

- `ggplot` and `plotly` for visualization

Once data is read in from CSV to dataframe (one file per metric), they are filtered to contain only the information described above. Then they are all merged using `merge` based on country code and year. From here, any NA are dropped.

For the prediction problem, the target is next year's inflation number for each country. This data is first extracted from consolidated data to update the year and then merged back as target. Finally, the train and test data are split in 70% and 30%, respectively, in chronological order i.e. the test data occurs strictly after the training data by country.

Feature Overview

The core economic indicators we have include gross GDP, GDP per capita, and GDP growth showing economic output. Population, specifically % of population that is of working age along with employment rate shows the health of the labour market. The current inflation, household savings and spending, and rent and share prices illustrate the consumption and price levels. Finally, interest rate, reflects the government's action to control and drive the economy. When the economy is heating up (inflation is higher), governments tend to raise interest rate to cool off the market. Conversely, when the economy is cooling down, interest rate is decreased to act as stimulus to drive the economy.

Results

Looking at the correlation between inflation and key economic indicators, we see that there is a high correlation between inflation and interest rate which is expected. Another metric that has a more significant correlation with inflation is rent growth, especially in top 5 countries with most data samples shown in below table. This is not surprising since rent growth directly reflects increase in cost of living.

Table 2: Correlation between Inflation and Key Metrics

Entries	Cor.GDP	Cor.HouseCost	Cor.SharePx	Cor.EmployRate	Cor.InterestRate
748	0.231	0.615	0.018	-0.195	0.757

Table 3: Corr. b/w Inflation and Key Metrics – Top 5 Countries

Country.Code	Entries	Cor.GDP	Cor.HouseCost	Cor.SharePx	Cor.EmployRate	Cor.InterestRate
USA	50	0.680	0.915	-0.090	-0.409	0.782
AUS	42	0.352	0.835	0.266	-0.605	0.857
CAN	40	0.414	0.861	-0.100	-0.489	0.847
NZL	35	-0.217	0.791	0.180	-0.085	0.853
GBR	26	0.093	0.256	0.251	-0.469	0.134

Referring to the interactive charts of inflation, interest rate, and GDP growth, we see that inflation and interest rate follow very similar trajectory. From the historical data, there are some interesting notes

- Inflation, and in turn interest rate, are quite high in 1970-80s
- In 2008, almost all countries have negative GDP growth because of the global financial crisis. Interest rate has also taken a deep dive at the time
- In 2014, Ireland's GDP growth stands out from all other countries. This was because they lowered corporation tax, allowing large multinational corporations to relocate their economic activities, therefore

driving the economy. Consequently, GDP growth in 2015 went back down.
- Inflation and interest rate are extremely low in 2020 due to COVID-19 pandemic and governments trying to stimulate the economy. In some cases, we can see negative inflation and interest rate. On the other hand, GDP growth was negative for many country.

Predictive Models

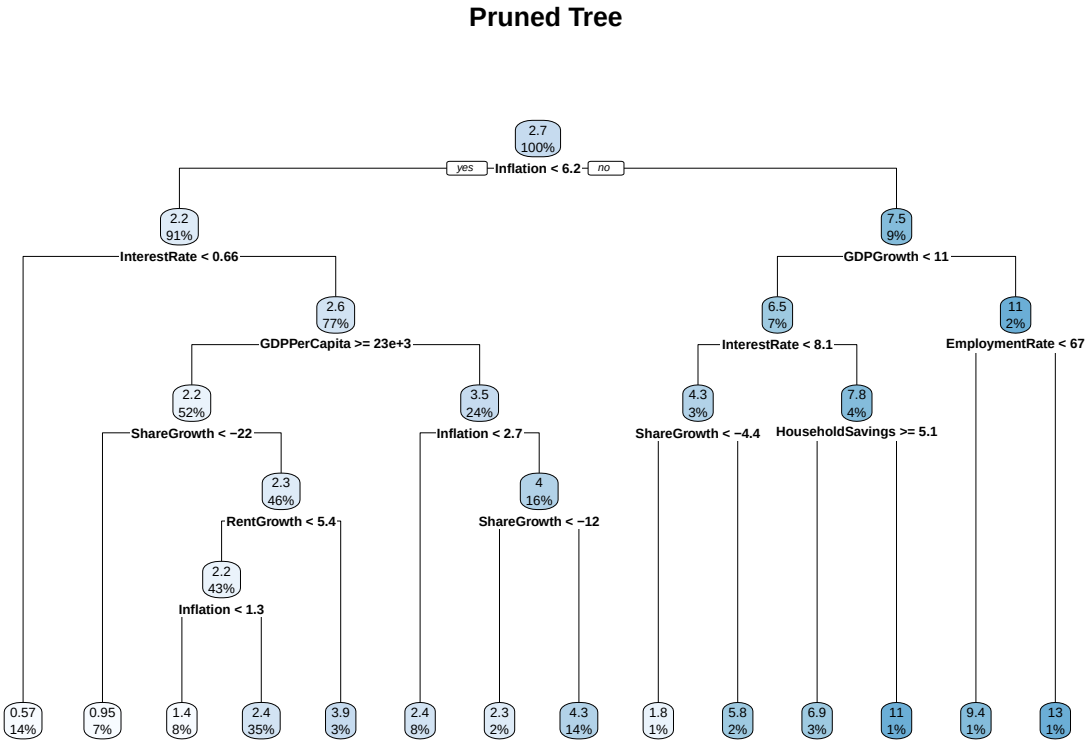
In order to build a model that predicts each country's inflation for the following year given current state of the economy, all the features that we have in the data are used, including current year's inflation number. The models that are used are

- Pruned Decision Tree
- Bagging
- Random Forest
- GBM
- XGBoost

We can see from the variable importance plots that different models give quite different results. However, the common theme between all models is that current year's inflation is the most important feature.

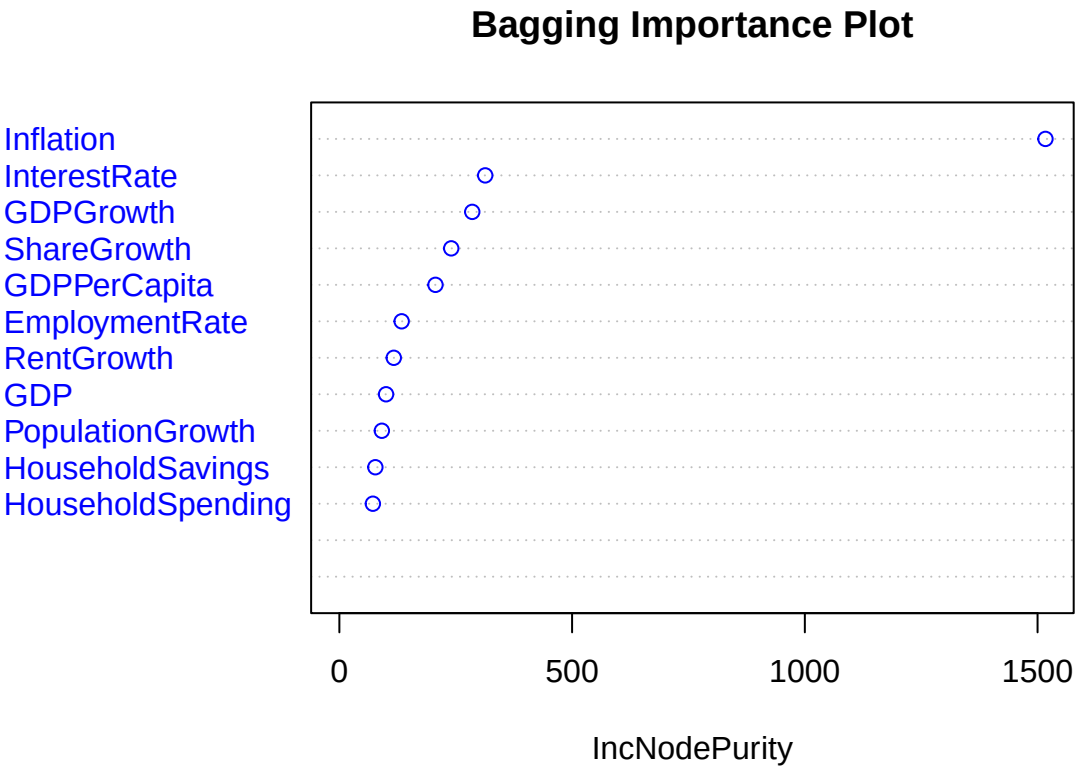
Decision Tree

We can see that the pruned tree is quite complex, it used various number of features. The main ones are inflation, GDP growth, and interest rate.



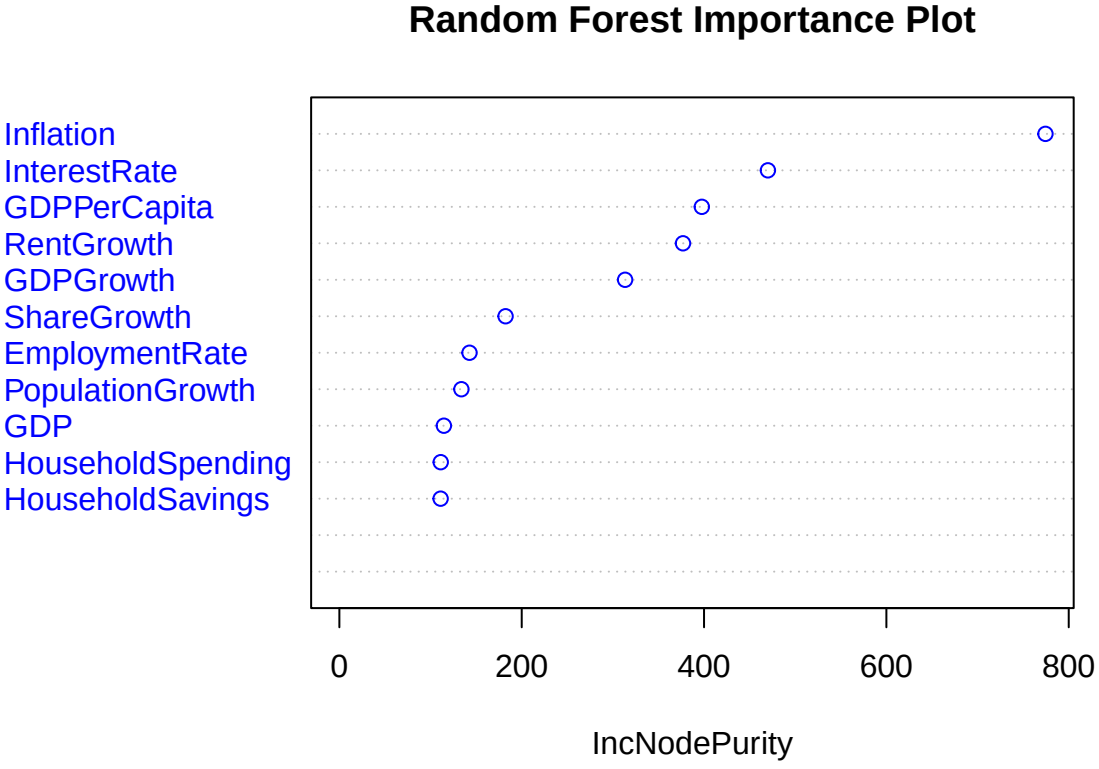
Bagging

For bagging, we set the number of variables sampled as candidates to be the same as the number of variables, 11. We can see that inflation is the most important feature.



Random Forest

From the plot, inflation is still the most important feature, but some other features such as interest rate and GDP per capita has became more important.

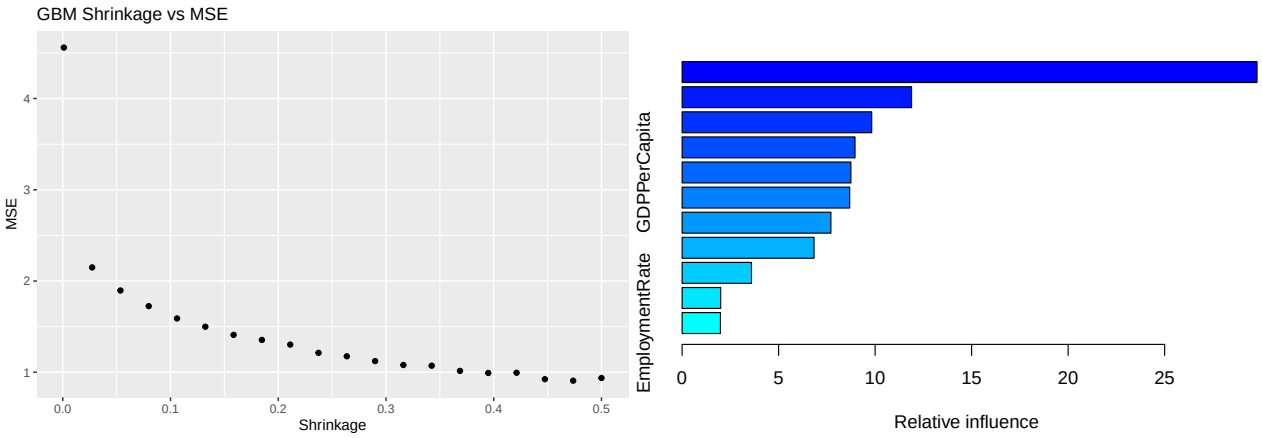


GBM

In GBM model, we try 20 different values of shrinkage from 0.001 to 0.5. By comparing the MSE (mean squared error) at different shrinkage value, we identified that shrinkage of 0.5 is best suited for the GBM model.

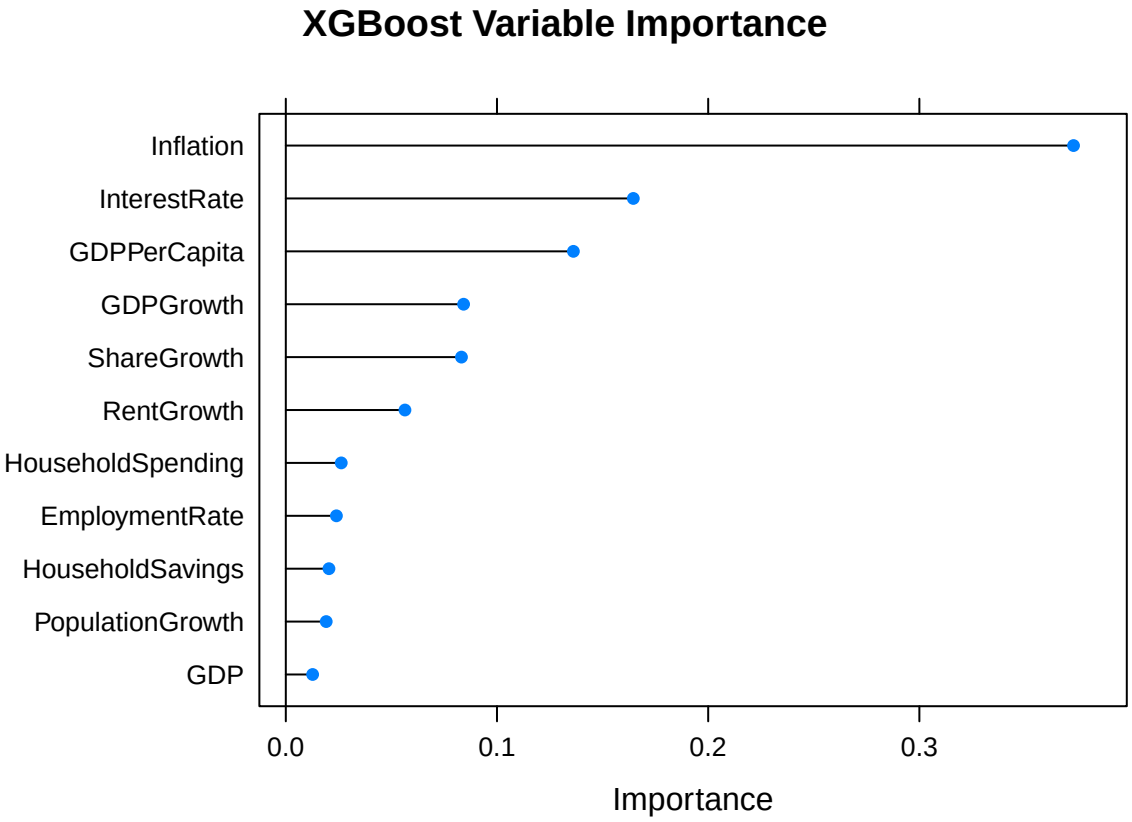
Table 4: Variable Importance

Feature	Relative.Influence
Inflation	29.79
RentGrowth	11.89
GDPGrowth	9.82
PopulationGrowth	8.96
GDPPerCapita	8.74
InterestRate	8.68
GDP	7.71
ShareGrowth	6.83
HouseholdSavings	3.60
HouseholdSpending	2.00
EmploymentRate	1.98



XGBoost

For XGBoost, we perform a grid search on η, for 20 values between 0.001 and 0.5.



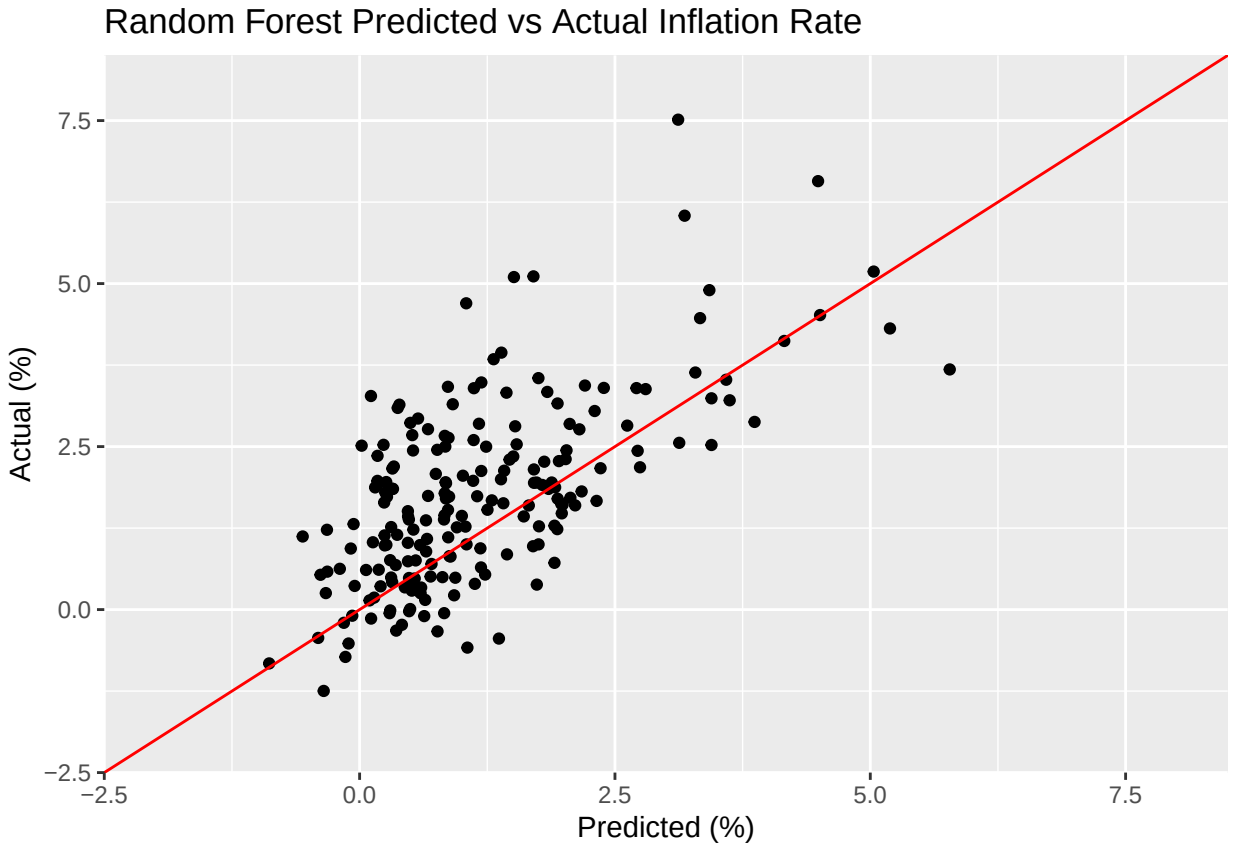
Model Performance

Based on the model performances, we can see that Random Forest performs the best out of all the models. All the models' MSE are actually quite close to each other since they all identified the most relevant features related to future inflation. Boosting is the exception where it's performance is quite different from other models; from the importance chart we see that interest rate is not important in the boosting model whereas in other models, it is much more important. This sets boosting apart from other models.

Table 5: Methods and RMSE

method	RMSE
Regression Tree	1.8685
Bagging	1.7943
Random Forest	1.5276
Boosting	3.3405
XGBoost	1.6803

The following graph shows the performance of the Random Forst by plotting the predicted inflation rate against the actual inflation rate. The red line is the 45-degree line which indicates the case of perfect prediction where all points should lie on the line. From the graph, we can see there is a linear relationship between the predicted and actual inflation. The implication is that the Random Forest model is capable to predict future inflation to some extent. This is important because it can help policy makers to better understand the future state of economy and adjust policies accordingly.



Conclusions and Summary

In summary, through this analysis on inflation, we constructed a model that predicts next year's inflation given current state of a country's economy. We also got a better understanding of how inflation is tied to or affects the economy. Out of the key economic indicators we used, current inflation, interest rate, and rent growth turned out to be the most relevant features. Multiple machine learning models were used for the study and random forest has the best performance in terms of MSE even though model performances were similar.

By understanding inflation and being able to predict it, it allows economists and policy makers to better manage the state of the economy using macro and micro-economic tools.