Capstone 2: GDP Forecast

Problem statement formation

What investment opportunities exist for portfolio managers and investors to increase their return, either by investing into a higher growth country, or adjusting their portfolio exposure globally based on the country growth forecast?

Context

Portfolio managers and investors look for global investment opportunities. But uncertainties on a country’s growth may affect their investment decision. To have better insights into a country’s growth will remove one barrier from their process, which is the goal of this project.

Criteria for success

In order to succeed, a country’s GDP growth will be forecasted with as little error as possible in the next 24 months based on 6 factors. Unemployment rate, consumer price index (CPI), domestic market manufacturing producer price index (PPI), housing prices, household debt to net disposable income, and government spending to GDP.

Scope of solution space

The higher the GDP growth of a country, the more the opportunities will be from an investment standpoint. Vice versa, lower GDP growth will lead to more cautions for pouring investment into the country.

Constraints

Unexpected stochastic events are not predictable. They can be an environmental catastrophe, a sudden market crash, a pandemic outbreak, or any other bubble burst.

Stakeholders

Portfolio managers or investors who take global economy as a consideration in their investment decision making process are my clients, as this project will focus on how predicting global GDP change in the next 24 months. Investment opportunities arise in countries with higher GDP growth.

Data sources

OECD.org

DSM steps

Step 1 Business understanding

We are trying to solve the problem of forecasting GDP growth of a country through different variables. This will help us understand two main keys:

1. Which variables correlate the most to GDP growth?
2. Are there any signs that would tell us if a market crash is underway?

Step 2 Data wrangling

Data collected may contain missing values and different frequency, e.g., monthly, quarterly, annual. We will clean and select appropriate data for each time-series to perform multivariate time-series analysis

Step 3 EDA

After merging all time-series into one dataframe, we will further impute any missing values in the resampled dataset. We will also visualize the graph and test their stationarities with Dickey-Fuller and KPSS tests. Adjustment like differencing, log or square root transformation may be needed accordingly.

Step 4 Preprocessing and training

We will split the data into training and test sets. Since this is a multivariate analysis, we will use Vector Autoregression (VAR) instead of the regular ARIMA model to train our time-series data.

Step 5 Modeling

Once the model is trained, we will invert the time-series back to its original form if any differencing or transformation is used in the stationarity tests.

Step 6 Documentation

A slide deck and report will be written to communicate the key takeaway of this project.