Unemployment Rate Forecasting using Machine Learning

Susan Jiao
jjiao25@wisc.edu

Yi Xiao
yxiao84@wisc.edu

Yuanhang Wang
wang2243@wisc.edu

1. Introduction

Policy researchers conduct extensive economic forecasting to make good suggestions to policy makers. However, forecasting economic indicators is challenging, since indicators can be influenced by a wide-range of factors such as politics and financial markets. The task has been especially difficult in today's highly internationalized and dynamic world. Unemployment rate is one of the most common and important economic indicators. Unemployment rate is a significant factor in many government decision making processes, such as setting interest rates and designing social welfare programs, and those policy decisions often have direct impacts on people's daily life. Thus building accurate forecasting models of unemployment rate and other important economic indicators is an active research area for many macroeconomists.

The most standard methods of time-series forecasting are moving-average models (MA) and autoregressive (AR) models which predict future observations using recent observations [3]. Some commonly used ones are threshold autorgressive (TAR) models, Markov switching autoregressive (MSA) models, linear univariate autoregressive integreated moving average (ARIMA) models, and bivaritae vector autrogressive moving average (VARMA) models[5]. Those models are widely used in macroeconomic indicator forecasts, including unemployment rate. Another popular approach to unemployment forecasting is a simple random walk model. It forecasts tomorrow's unemployment rate using on today's unemployment rate and an additional random shock[3]. Besides statistical models, consensus forecasting is also a widely-used methodology. Instead of using certain specific models, it combines different kinds of models that draws from a variety of techniques and methods, ranging from econometrics to meteorology[3].

Those traditionally used methods for unconditional forecasting are mostly linear models that have high data demands. More recently, researchers in economics and political science began to utilize deep learning techniques, particularly neural networks such as Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU), to make predictions. Researchers in economics found deep neural networks achieve higher accuracy with limited data than transitional econometric methods do. Cook and Hall (2017) predicted civilian unemployment using models that built upon four different neural network architectures, and those models outperformed consensus forecasting at shor time horizons. In particular, they found that an Encoder Decoder architecture outperforms baseline models at every forecast horizon[6].

In this project, we will forecast unemployment rate using machine learning techniques such as random forests and neural networks and will use standard regression methods as benchmark models.

2. Motivation

Unemployment rate attracts great attention from the public section, the government section, and the business section, especially during recessions and political turbulence. Unemployment rate is a key indicator of a country's labor force market performance and reflects the general health of the economy. The Federal Reserve uses unemployment rate as an important factor in setting monetary policies, while investors utilize unemployment rate in sector-specific investment decisions. For those unemployed, social welfare programs offered by the government, such as unemployment benefits, have direct impact on their lives. Those programs are often designed by tracking changes in unemployment rates. Thus, accurate unemployment rate forecasts is important in promoting economic stability, social welfare of the people, as well as business investments.

From an individual perspective, when workers are unemployed, their families experience great financial losses and decrease in purchasing power. The decrease in demand for goods and services could lead to financial losses in more families and negatively affect employees in other fields. [2] In addition, unemployed people might experience mental health problems, such as loss of self identity, and even physical health problems. Many unemployed people suffer from social isolation. If no social support is provided, the situation would easily deteriorate, and in severe cases, it might even lead to social disruption. Therefore, the role of government in providing assistance to those people is vital to the health of those people and families as well as the general stability of a society.

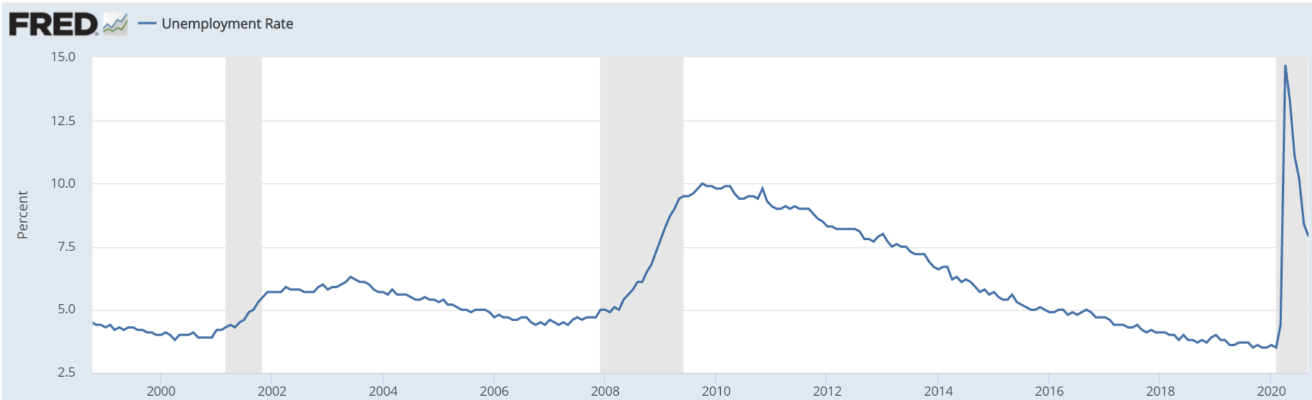

Figure 1. twenty years of monthly unemployment rate data from 2000 to 2020. [1]

Overall, understanding the trend of unemployment rate is important for the health of an national's overall economy and for the health of each individual. Building a reasonable and feasible forecasting model of unemployment rate will help policymakers make better decisions to stabilize the economy and ensure the benefits and life qualities of citizens.

3. Evaluation

We want to predict the unemployment rate for a certain period. There is not a "best" criterion for measuring the forecasting accuracy, but for simplicity, we will only consider two common criteria: Root Mean Squared Error and Mean Absolute Error. We will divide the dataset into training, validation, and test sets. We will use the training set to train our models, the validation set to tune parameters and evaluate the performance on the test set. The average RMSE and MAE on the test set will be a good representation of the prediction power of our model.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \quad (1)$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y}_t| \quad (2)$$

As we are dealing with time-series data, we will compare our models when using different time-horizon features. For example, we will look at the performance of each model when using only the previous 3 months as features, 6 months as features, and 1 year as features. We will find the best model in each setting.

We will use a simple linear regression model as a baseline model, and compare the performance of other models to it. We consider our project to be successful if our model outperforms the baseline model.

Our data only include several economic indicators, including unemployment rate. Note that there are many other factors that affect unemployment rate, which our model will not account for. For example, the recent pandemic crisis had caused the unemployment rate to spike. Such factors do not have a clear patter, so it is hard to incorporate them in our model. We may add new predictors if we found useful information as we research through this area.

4. Resources

We obtained our data from Federal Reserve Economic Data [4]. The dataset contains different economic indicators recorded monthly from Jan. 1959 through Oct. 2020, including the unemployment rate. We will select and use some of the indicators as additional predictors.

Each group member will primarily use his/her laptop to conduct tasks. The computation tasks will mainly be done on Jupyter notebooks using Python packages such as Numpy as Scikit-learn. In addition, as we might use deep learning models, we may also use Google Colab as it provides faster computation with GPUs.

5. Contributions

Yuanhang will be responsible for data pre-processing, feature selection, and generating the baseline model. Yi will develop machine learning models. Susan will run the deep learning neural network models.

All team members will write the report together.

References

[1] "Unemployment Rate (UNRATE)". *Federal Reserave Bank of St. Louis*.

[2] Canberra. Inquiry into issues specific to older workers seeking employment, or establishing a business, following unemployment. *Commonwealth of Australia*, 2000.

[3] A. S. Hall. Machine learning approaches to macroeconomic forecasting. *The Federal Reserve Bank of Kansas City Economic Review*, 103(63):2, 2018.

[4] M. W. McCracken. ”FRED-MD and FRED-QD: Monthly and Quarterly Databases for Macroeconomic Research”. *Federal Reserave Bank of St. Louis*.

[5] A. L. Montgomery, V. Zarnowitz, R. S. Tsay, and G. C. Tiao. Forecasting the us unemployment rate. *Journal of the American Statistical Association*, 93(442):478–493, 1998.

[6] A. Smalter Hall and T. R. Cook. Macroeconomic indicator forecasting with deep neural networks. *Federal Reserve Bank of Kansas City Working Paper*, (17-11), 2017.