Capstone Machine Learning

Connor Vaughan

The primary use of machine learning in my project will be for forecasting future unemployment rates by training a model with data on past unemployment rates. The primary tool used to do this was the ARIMA forecasting package. This package was chosen because the data is seasonal and has linear trends which allows ARIMA to create forecast models. The data from 1973 to 2017 will be used to train the model. Then the model will predict future values of the unemployment rate for each month in 2018, then these values will be compared to the actual unemployment values for each month in 2018 to see how close the model got. The closer the model’s predictions are to the actual values, the more accurate the model is at forecasting. There are also statistical measures used to determine the accuracy of forecasting models, particularly MAPE and RMSE. MAPE is an acronym for mean absolute percent error, and it is found by taking the difference of the actual and the forecast then dividing by the actual and multiplying by 100, then take the mean of all those values. RMSE is an acronym for root mean square error and is found by taking the standard deviation of all the errors between the forecast and the actual values. Both of these tools give an empirical measure on the quality of my machine learning model.