**Model Summary**

**Neural Network**

A test set of voter data was used to predict the number of voters for 2020, in the states of Arizona, Florida, Michigan, North Carolina, Pennsylvania and Wisconsin.

**Pre-Processing**

The data was gathered from the Federal Election Committee (FEC) and placed in a table, in SQL, called agg\_county\_votes. Then, the data types and null values were identified and updated. The numeric and categorical columns were listed, then merged with getdummies, to create the X value.

**Feature Engineering**

The features are listed as follows:

* blue\_votes Democratic votes
* red\_votes Republican votes
* other\_votes All other votes
* election\_year Presidential election years (2000, 2004, 2008, 2012, 2016)
* PopPct\_Urban Percentage of the county’s urban population
* Unemployment County’s unemployment rate
* PopDen\_Urban Density of the county’s urban population
* PopPct\_Rural Percentage of the county’s rural population
* PopDen\_Rural Density of the county’s urban population

**Training / Testing Split**

The data was split to train on 75% of the data and test on 25%, using a random state of 78. The data was then scaled, using MinMaxScaler.

**Model Choice**

A Sequential model was used, with Relu and Linear dense keras layers. Then, the data was optimized by using the Adam and the Mean Squared Error loss compiling method. To fit the model, it was run with 100 epochs.

**Model Accuracy**

The r2 score on the training data was .025555 and the testing data was .000937. However, running a deep model produced greater results, train was .994397 and test was .988827. It is likely that this is due to overfitting.

**Model Success**

The model was not successful so we looked at Linear Regression and Unsupervised Learning Models instead.