# Insights Report

Wrangle and Analyze Data Project Submission
Data Analyst Nanodegree Program
Jessica Ertel

I created two sets of visualizations from the wrangled data. The first visualization was a set of bar charts using the image predictions dataset. I wanted to investigate the most popular breeds of dogs from the WeRateDogs archive. According to the neural net's predictions, it appears the two most common breeds appearing in images are golden retrievers and Labrador retrievers. An example of the chart illustrating the neural network's first prediction is below.

The Chihuahua is also a popular breed, ranking in the top four across all three prediction attempts. It's important to note two items when interpreting this chart. The first is that a single tweet can have up to four images, which means that a single tweet of a Golden Retriever might be identified as a Golden Retriever four times by the neural net. The second is that this chart does not account for how confident the algorithm is in its prediction, which means it could be predicting with a very low confidence.
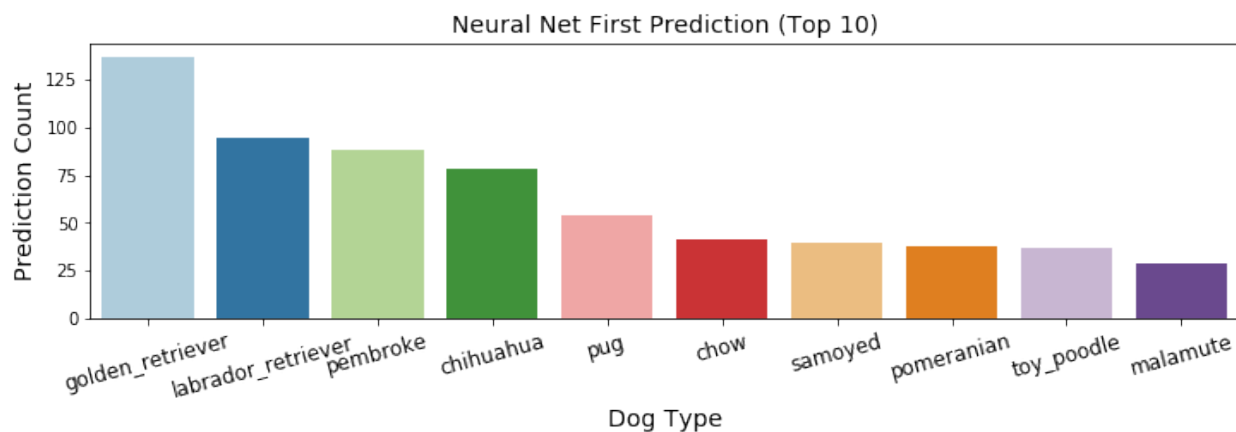
The second visualization I created is a scatter plot using the data from the enhanced twitter archive. I was interested in seeing if there were any patterns in the popularity of posts over time. According to the chart we see a positive correlation between the number of 'favorites' or likes and time. In 2017, tweets appear to be getting a higher average number of likes than in the beginning of 2016, which sees a concentration around the 5,000 mark. It's possible this increase over time is attributable to the WeRateDogs twitter account seeing an overall increase in popularity.

Tweet Popularity