Insights from analysis of WeRateDogs Twitter Data

By Kyle Wilbert

Overview

For this Udacity project, we were asked to wrangle data about the Twitter account WeRateDogs (@dog_rates).

The first dataset was an archive of tweets from WeRateDogs, presented to us as a CSV.

The second was a TSV of image predictions results that we downloaded programmatically. The project creator passed the photos from the tweets into a machine learning project and it spit out these predictions.

And the third source we scraped from Twitter using the Tweepy Python library.

After wrangling and cleaning the data, I was left with 1992 tweets (no retweets).
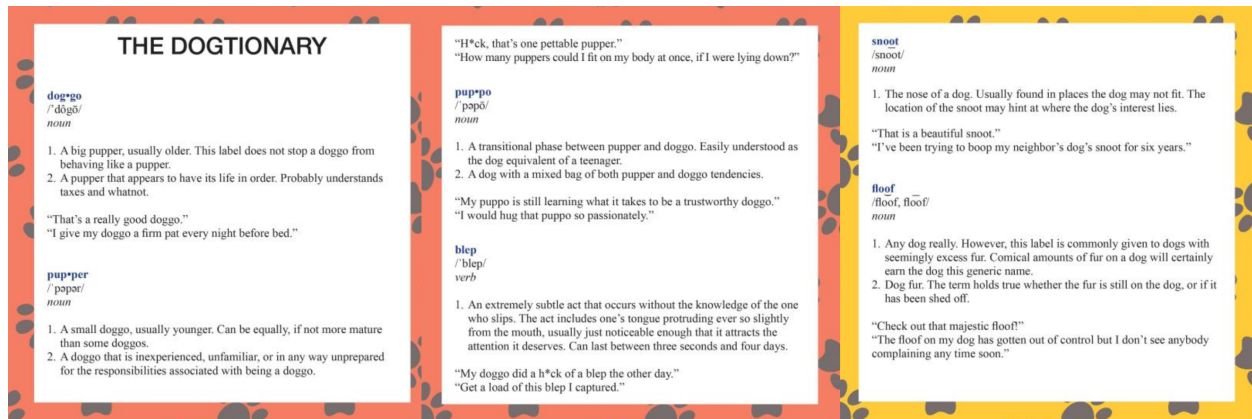
Sources of Tweets

The tweets came from three different sources – Twitter for iPhone, Twitter Web Client, and TweetDeck.
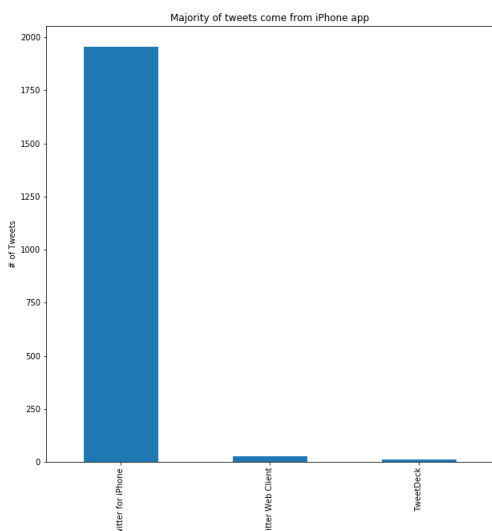
- 1,954 tweets were from Twitter for iPhone (98.1%)
- 28 tweets were from Twitter Web Client (1.4%)
- 10 tweets were from TweetDeck (0.5%)

Distribution of Dog Stages

Many of the tweets come with WeRateDogs special dog stage.

In our dataset, we had ratings for doggo, pupper, puppo, and floofer

- 306 of 1992 (15.36%) tweets had at least one dog stage
- 74 of 1992 (3.71%) tweets were doggo
- 8 of 1992 (0.40%) tweets were floofer
- 212 of 1992 (10.64%) tweets were pupper
- 23 of 1992 (1.15%) tweets were puppo

Image Prediction Accuracy

Of the 1,992 tweets, 1,684 had image predictions where at least one of the image predictions was in fact a dog (sometimes the algorithm identified other objects in the image, like an orange or a shopping cart).

In the dataset, we were given three predictions for each tweet. Here are some informal charts showing the distribution of confidence scores for each prediction. For each, the X-axis is the level of confidence for the prediction (%) and the Y-axis is the number of tweets in that bin.

(cont'd)

Distribution of Confidence Levels for Prediction 1

Distribution of Confidence Levels for Prediction 2

Distribution of Confidence Levels for Prediction 3
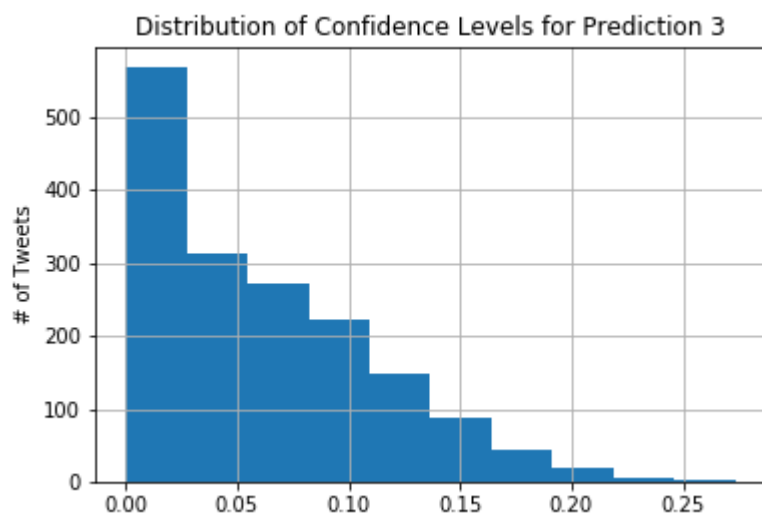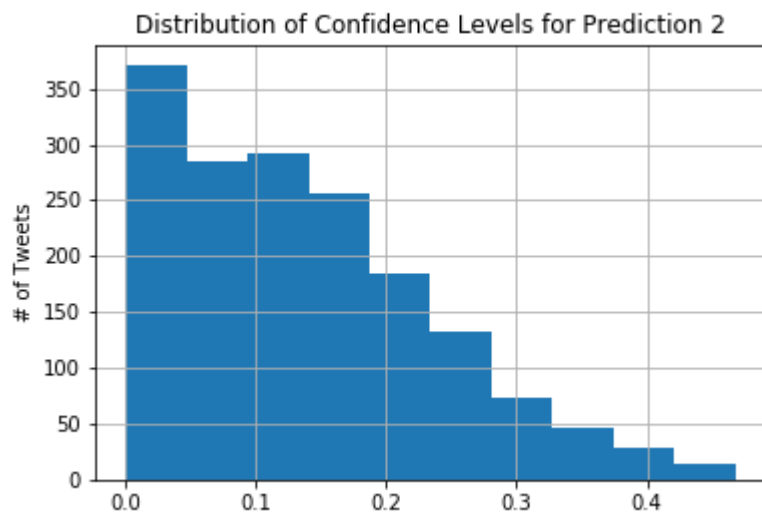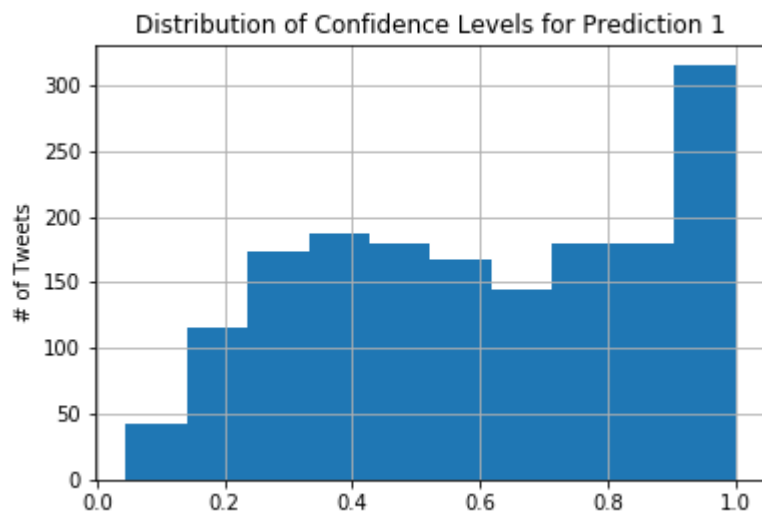
The first prediction had the widest distribution of confidence, ranging from .044 to .999, with a mean confidence of .605. So the image prediction is somewhat successful, but not high enough to rely on for any serious study.

The other two predictions were less successful.
- Prediction 2: range 0 – 0.468, mean 0.137
- Prediction 3: range 0 – 0.273, mean 0.061

Based on this brief descriptive analysis, I'd say the image prediction algorithm needs some more training and/or tweaking.