Overview

This bot generates tweets resembling a mix of Donald Trump and the Flat Earth Society.

I attempted to accomplish two things with the project:
1. Explore Natural Language Processing and Recurrent Neural Networks
2. Raise awareness about the similarities in language between Donald Trump and members of the Flat Earth Society

The publishing process was split into three phases: Generation, grammar ranking, and manual review.

Generation was done using Max Woolfe's textgenrnn library. The library allows for an LSTM to be trained on a text corpus, and text to be generated resembling that which it was trained on.

The ranking process involved taking a few thousand tweets and attempting to prioritize the ones that are most coherent.

Manual review involved me looking at the tweets in the order that the ranking process determined and choosing those fit for publishing.

Text Generation

Generation is done be training an LSTM model to predict a character given a prefix of characters. Once this relationship is known, a seed string has its most likely character predicted and appended to the prefix string. This process is repeated to produce text.

I initially wrote a set of tools to do the above process, but later realized that Max's library already had (almost) exactly what I was looking for. Hooray for open source!

The one tweak I made to the tools was adding the ability for specific characters to have the probabilities of their occurences manually adjusted. I used this to increase the chances of punctuations showing up, as the generations tended to frequently produce run-on sentences.

Grammar Ranking

Ranking is achieved by training a classifier to predict whether or not a tweet is grammatically correct, and sorting generated tweets by their inferenced grammar score.

Grammar information about a candidate tweet is produced by performing string transformations on the tweet and using the language_check API to register the errors in each transformation. String transformations include replacing @'s with proper nouns, removing hashtags, etc. The error counts for the transformations are then used as features to the grammar classifier. The

label is 1 if the tweet is really a twitter account, and 0 if the tweet was generated using random word permutations from the account. Classifiers for each of the sources (Donald Trump and all of the Flat Earth twitter pages) have their scores summed, and the ranking process orders them by total score.