Tweet Success Prediction

Bryson Oar, Ryan Outtrim

THE UNIVERSITY OF UTAH

Key Idea / Motivations

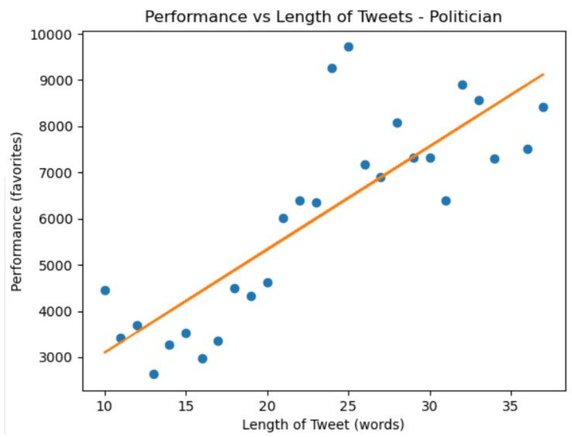
- Predict tweet success of a known Twitter account
- Find patterns to use for the model
- Create metrics to be used for other potential Twitter studies
- Give more insight to the underlying structure of a successful tweet that is inherent to most tweets

Data Set

- Used Twitter API to scrape data from a variety of Twitter users accounts
- Each user had all of their past tweets taken for content, date posted, likes, retweets and comments
- Our data looks at origin tweets, which are tweets that are not a reply to another tweet

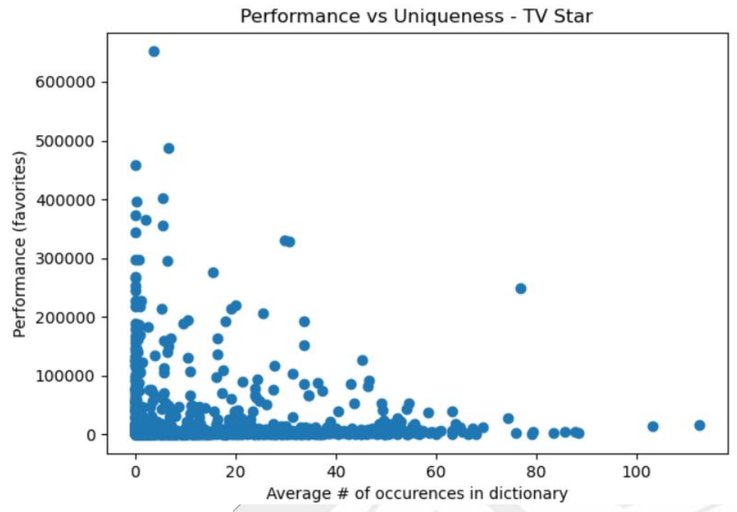
RESULTS / DISCUSSION

Length of Tweet



- Politicians, scientists, and news accounts all did best when their tweets had more words.
- TV stars and musicians did best when their tweets had less words.

Unique Vocabulary

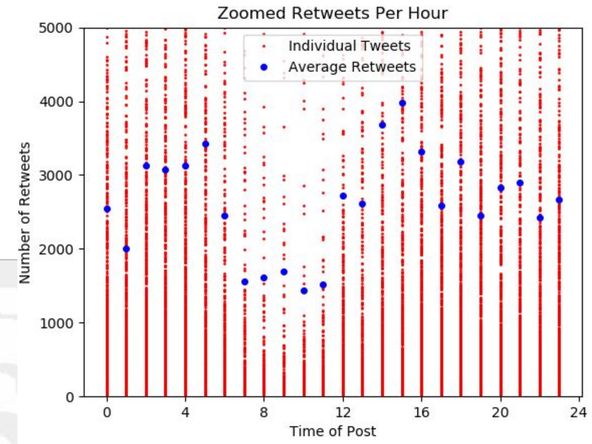


- Created a dictionary of words tweeter has used and graphed how unique the tweet is by perf.
- On average, the highest performing tweets tended to be more unique.
- No clear way to predict success from tweet uniqueness, though.

CONCLUSIONS

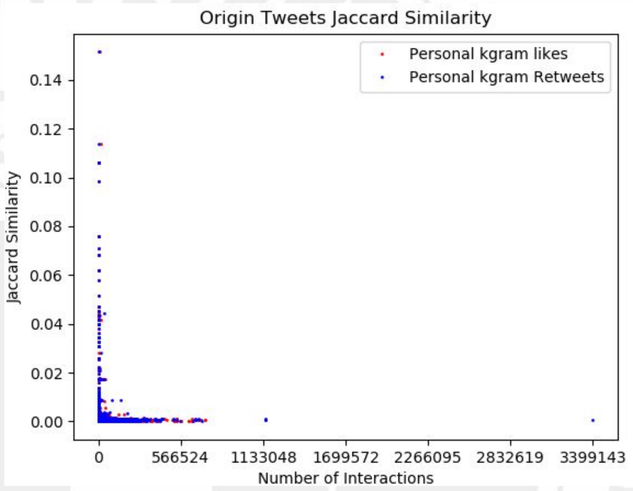
- We found that the most important parts of a tweet, other than the meaning itself, is its length, uniqueness, and time of day posted.
- Analyzing a tweet without knowing the meaning of it is difficult.
- Due to surprising results of some of the experiments, there were simply not enough parameters to create a multiple regression model with.

Time of Day of Post



- Look at retweets and likes of a post grouped by time of post sorted by hour
- High counts in morning and in the afternoon and evening

Comparing Similarity



- When a user says something unique to them it is more likely to perform well
- There is no relation to how a user talks when compared to the general Twitter populus