Modeling Tweet Sentiment

Charlotte Basch

Introduction

- Twitter has 166 million users
- 22% of American adults use twitter, with those users more likely to be more affluent and younger
- This makes Twitter an excellent resource for gauging consumer sentiment

•https://s22.q4cdn.com/826641620/files/doc_financials/2020/q1/Q1-2020-Earnings-Press-Release.pdf
•https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/

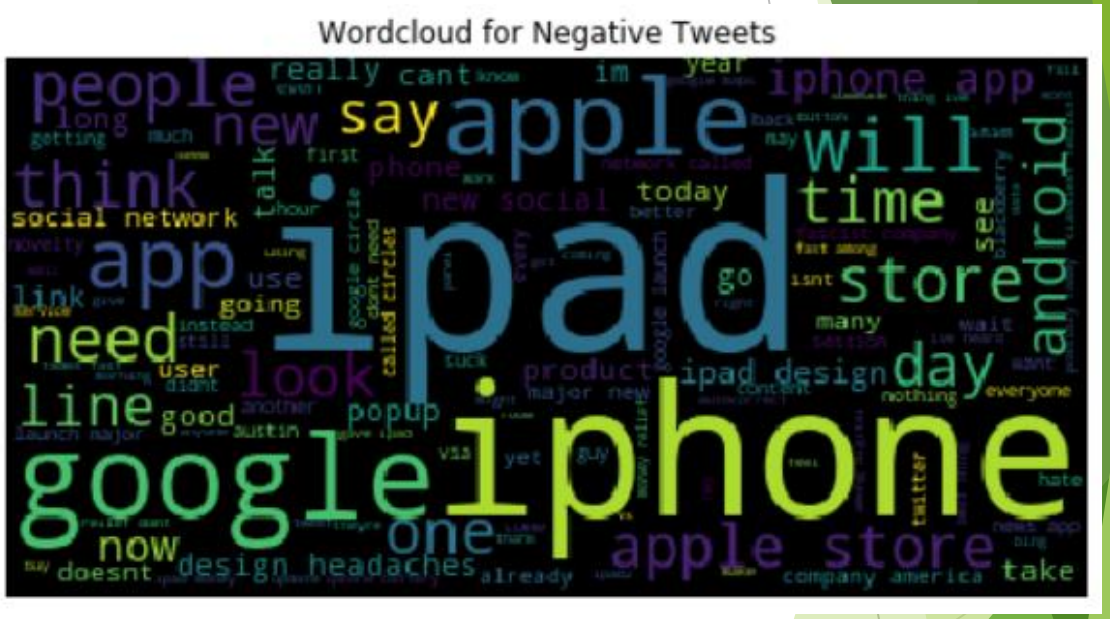
Data

- 9,000 tweets from CrowdFlower about Apple and Google products
- Rated as positive, negative, neutral, and unknown
- Used 3,500 positive and negative tweets

Wordclouds

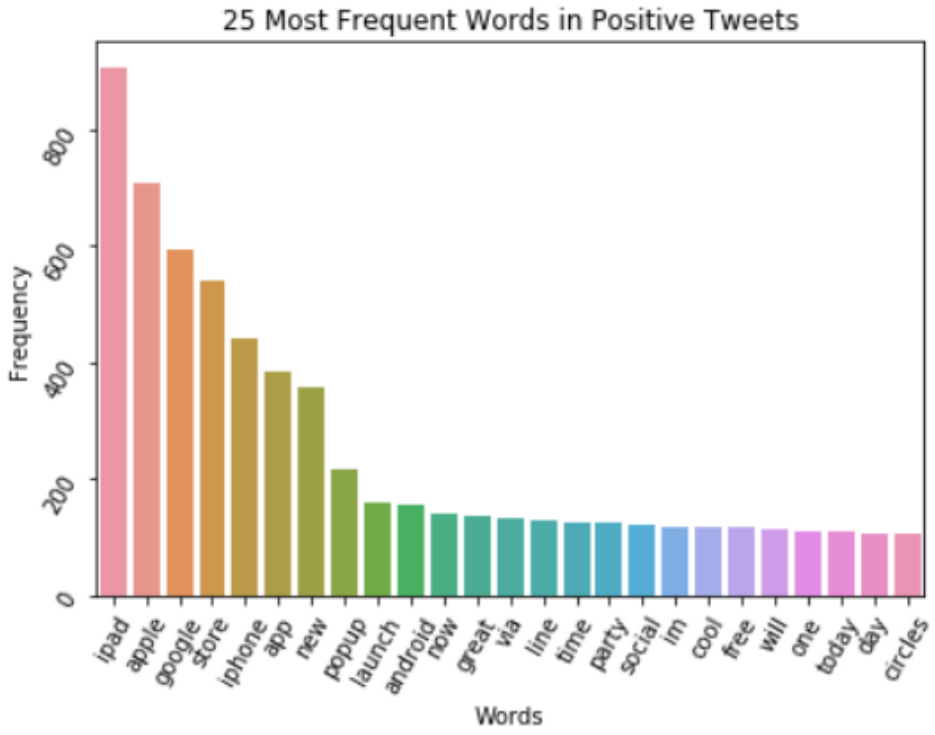
Wordcloud for Positive Tweets

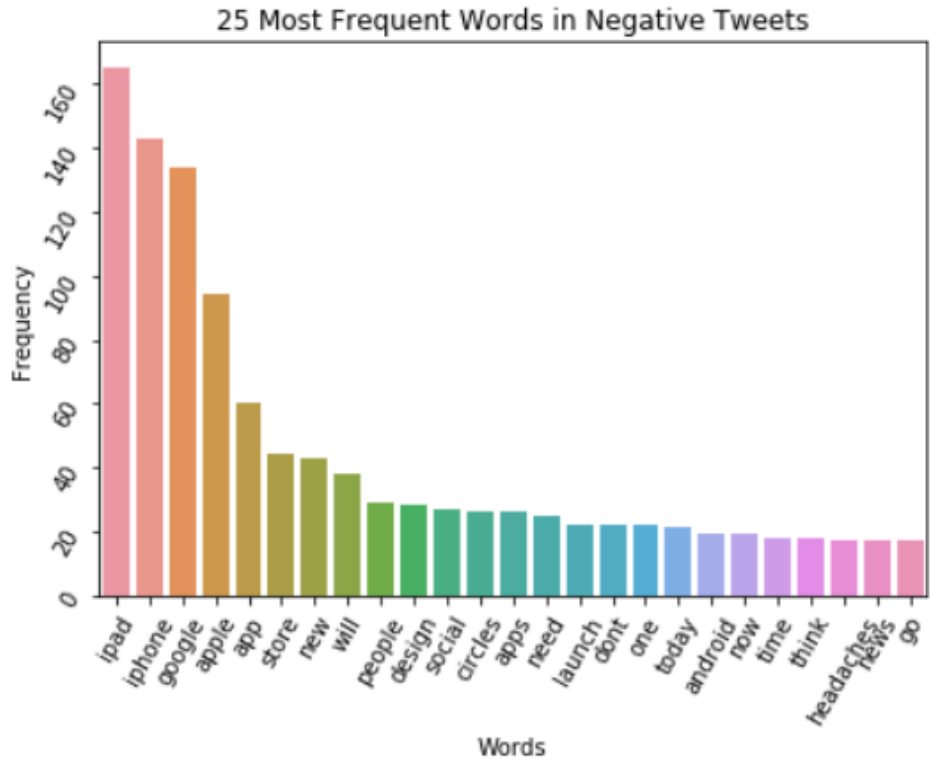
Wordcloud for Negative Tweets



- As expected, the company and product names show up frequently
- There are more positive words (i.e. great) in the positive tweets
- The positive tweets seem to reference new products

Most Common Words





- Interestingly, need only appears as a top word for negative tweets
- There are multiple references to new products (i.e. popup, party) in the positive tweets
- Overall the words are fairly similar

Modeling

- Long Short Term Memory neural networks
  - This is a network that is able to remember but also able to throw away information it does not need
- The best model was approximately 80% accurate

```
Model: "sequential_12"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_12 (Embedding)     (None, 32, 100)           275800
_________________________________________________________________
lstm_12 (LSTM)               (None, 100)               80400
_________________________________________________________________
dense_12 (Dense)             (None, 1)                 101
=================================================================
Total params: 356,301
Trainable params: 356,301
Non-trainable params: 0
_________________________________________________________________
```

Recommendations

- Having some kind of event around new products is related to positive tweets
- Look out for words that are talking about taking an action because it is related to negative tweets
- If you are looking for positive tweets about your product, the typical words like ‘great’ are in fact a big distinguisher between positive and negative tweets

Future Work

- Collect more data
- Add in neutral category
- Get data from a wider time period

Summary

- People tended to tweet more positively about the excitement of new products
- Action words, i.e. will, are more common in negative tweets, perhaps indicating that people are tweeting about their intention to stop using a product
- While the current model has fairly good accuracy, steps can be taken to improve the classification

Thank you!