Pay Attention to this Bag of Tweets!

Eric Adsetts and Eryk Wdowiak

Context

- How can brands analyze the way people feel about them by analyzing mentions on social media?
- We can use NLP to classify tweets by sentiment (Positive, Negative, Neutral)
- Used tweets about Apple and Google from the SXSW festival
- Able to classify tweets with an f1 score of .7

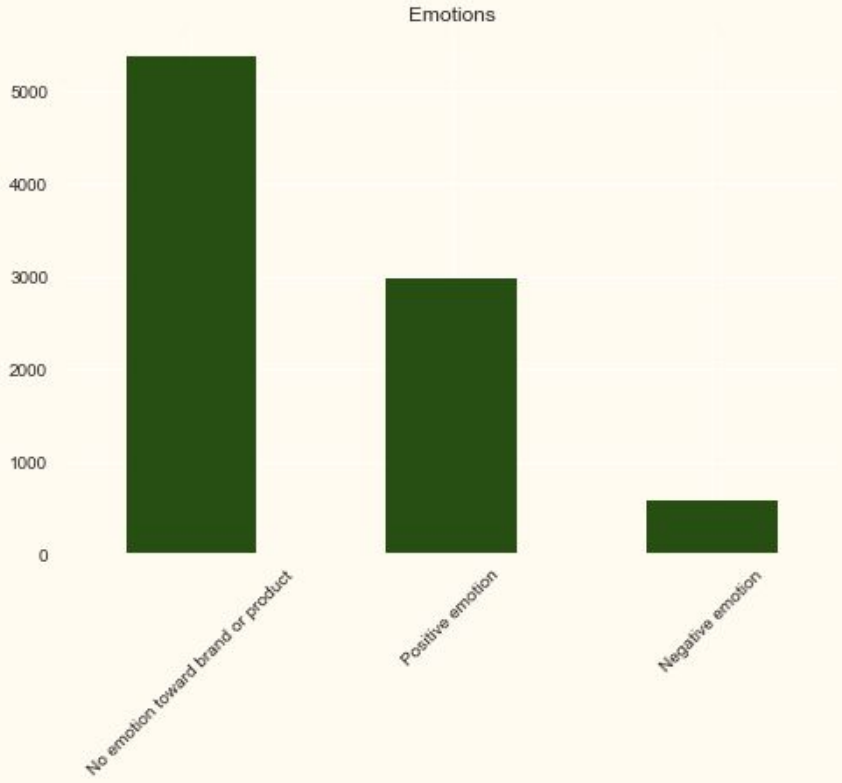
Our Data

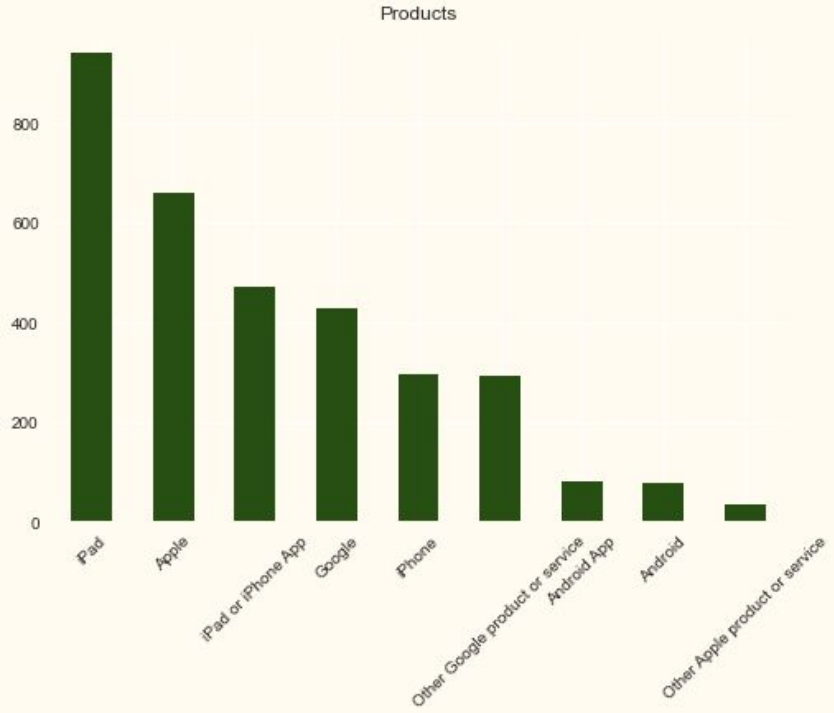
9092 Tweets

Mostly Neutral

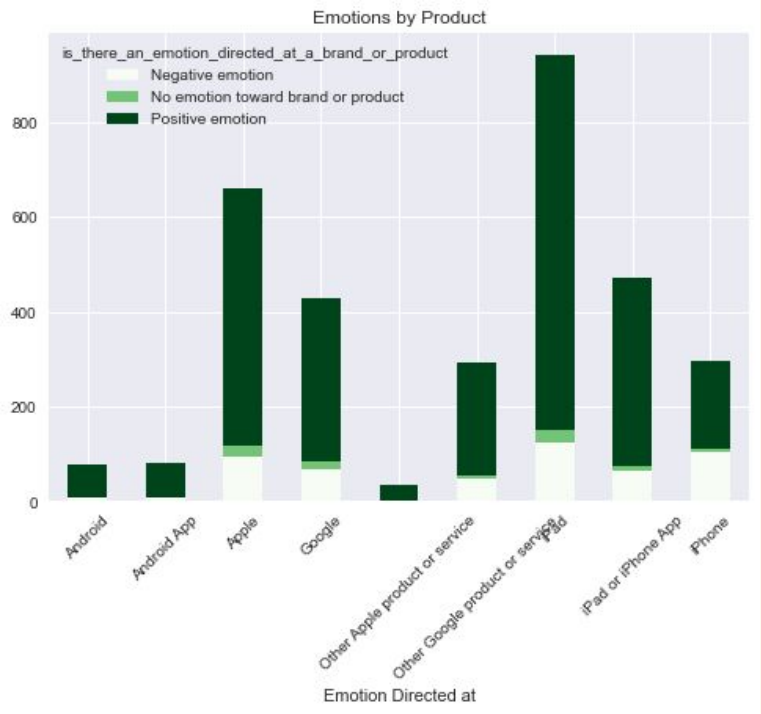
Mostly about Apple

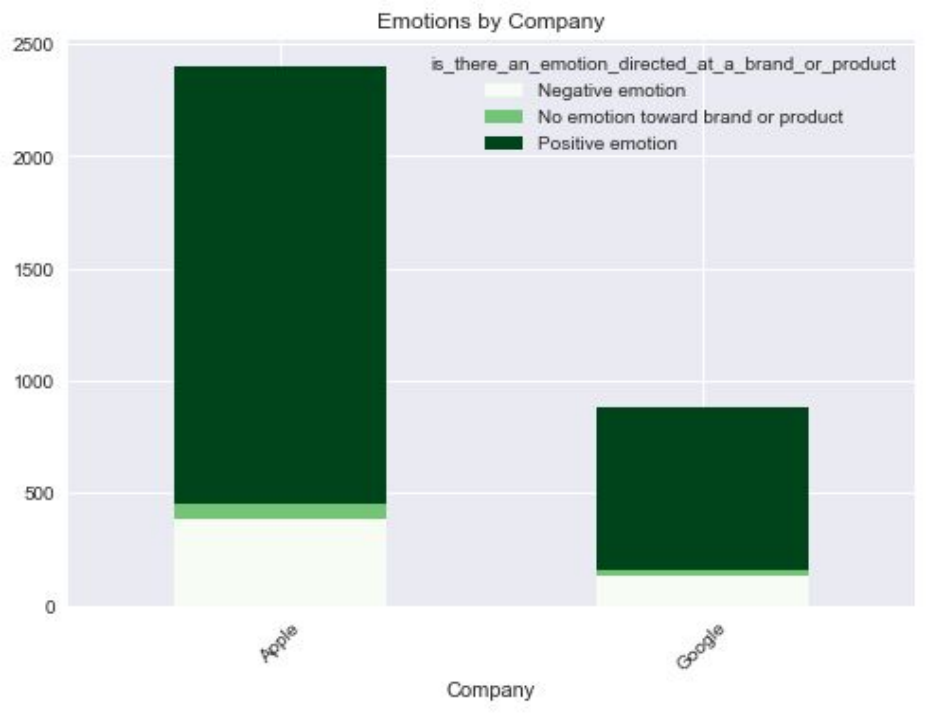
Our Data


Emotions
5000
4000
3000
2000
1000
0
No emotion toward brand or product
Positive emotion
Negative emotion


Products
800
600
400
200
0
iPad
Apple
iPad or iPhone App
Google
iPhone
Other Google product or service
Android App
Android
Other Apple product or service

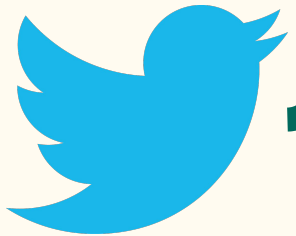
Our Data


Emotions by Product
is_there_an_emotion_directed_at_a_brand_or_product
Negative emotion
No emotion toward brand or product
Positive emotion
800
600
400
200
0
Android
Android App
Apple
Google
Other Apple product or service
Other Google product or service
iPad
iPad or iPhone App
iPhone
Emotion Directed at


Emotions by Company
is_there_an_emotion_directed_at_a_brand_or_product
Negative emotion
No emotion toward brand or product
Positive emotion
2500
2000
1500
1000
500
0
Apple
Google
Company

Process

Data from tweets about Apple and Google



Dating Cleaning and EDA

pandas

Visualization with Matplotlib and Seaborn

matplotlib

Lemming, vectorization and n-grams

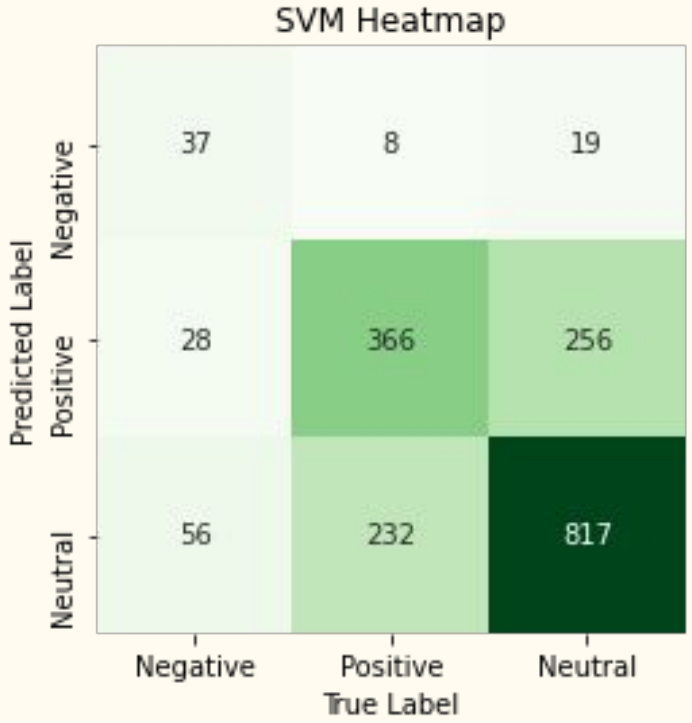
NLTK

Modeling

scikit learn

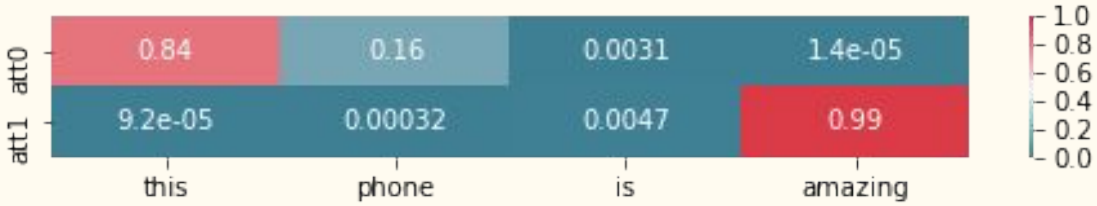
Results

Models	**F1 Score**
Naive Bayes	0.618
Random Forest	0.643
Logistic Regression	0.667
Decision Tree	0.598
K-Nearest Neighbors	0.609
Support Vector Machine	0.667
XG Boost	0.651
Voting Classifier	0.653



Self-Attentional Sentence Embeddings (1)

- Lin, Feng, et al. (2017) combine word embeddings and self-attention
- Word Embeddings
  - represent a word's semantic information in multidimensional space
  - semantic information can be used to classify sentiment
- Self-Attention
  - self-attention layers model the relationships between words in a sentence
  - those relationships can be used to identify the most important words



Self-Attentional Sentence Embeddings (2)

- If relationships among words were stationary, we could use this method to:
  - Classify
  - Understand that classification
- The trouble is that relationships among words is not stationary.
- So we will need more time to implement this approach.

Applications

- A company can use this information to measure sentiment towards their products and their competitors
- Next steps: further explore sentence embeddings to improve our model
- Could measure success of marketing campaign based on twitter sentiment