

# Sentiment Analysis of Tech Tweets

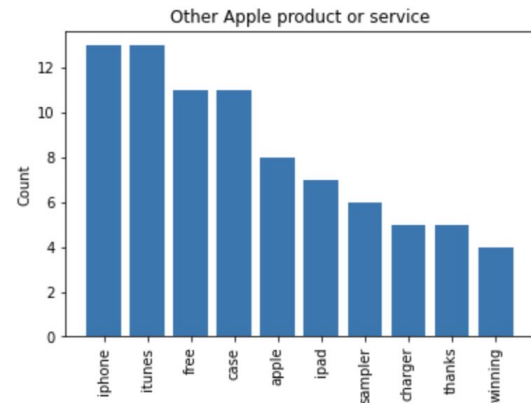
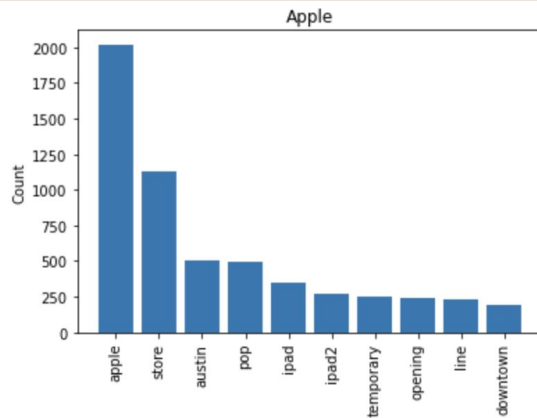
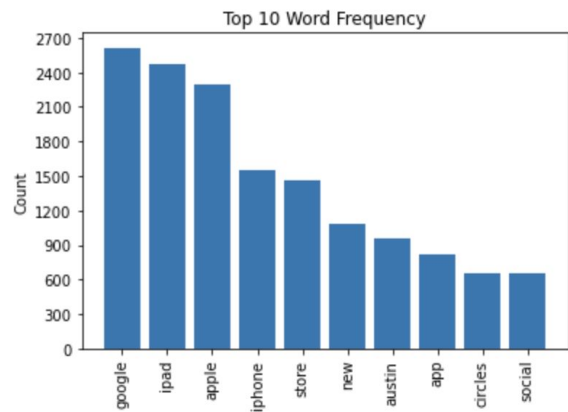
Cassidy Exum, Flatiron School

# Goals

- Investigate CrowdFlower dataset
- Refine a Sentiment Analysis Model
- Recommend marketing and product development opportunities to Apple

# About The Data

- 9093 rows
- 3 columns
- Date Added: August 30th 2013
- Frequency Distributions Below
- Large Class Imbalance



# Data Preprocessing and Feature Creation

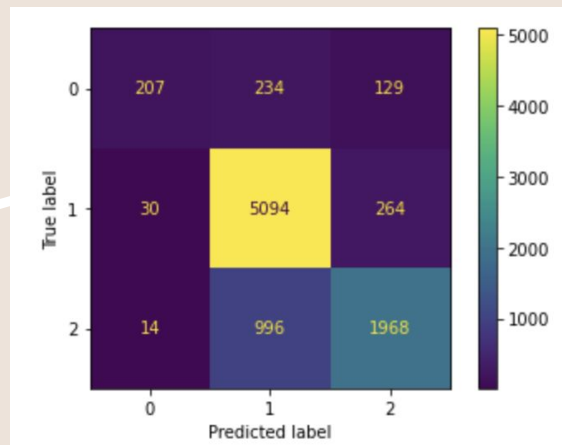
Different Models needed different preprocessing steps, some steps taken were;

- Removing stop words
- Stemming
- Finding the number of sentences in a tweet
- Determining if a tweet contained an emoticon
- Vectorization

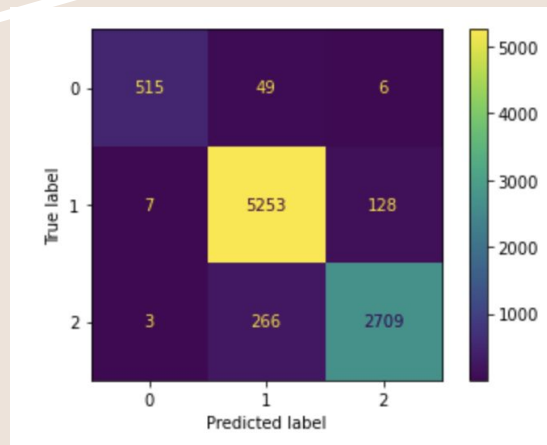
# Modeling and Evaluation

- First few models had training accuracy of about 60% and F1 score of about 58%
- Perceptron was the first success
  - F1 Score: 80%
- Random Forest model
  - F1 Score: 95%

Perceptron

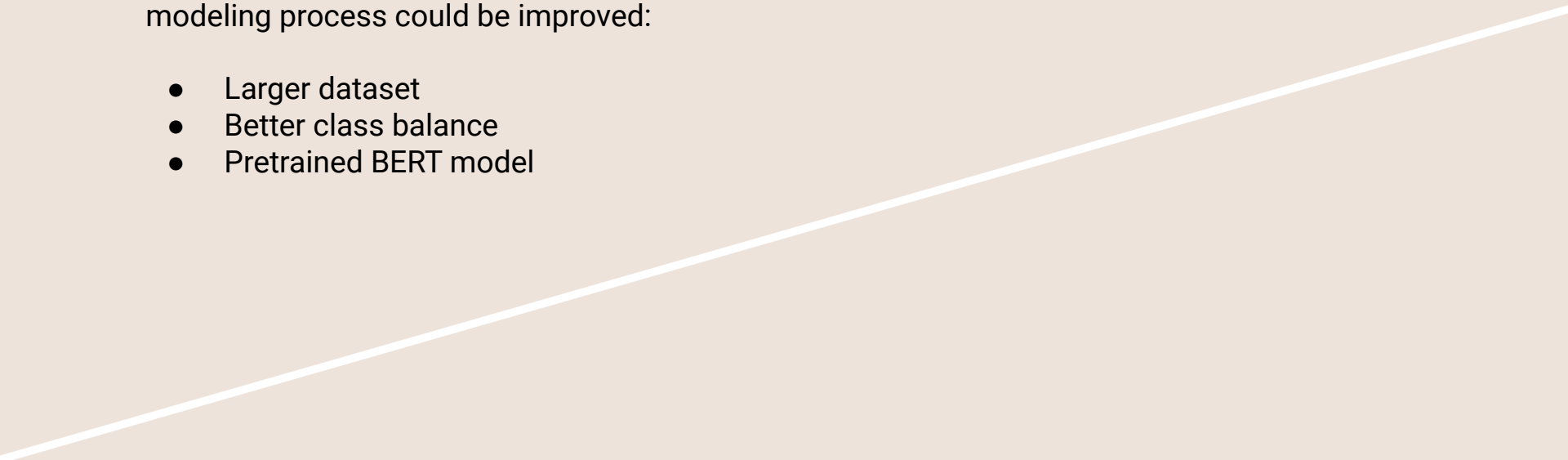


Random Forest  
(Best Model)



# How to Improve?

While I am happy with the score of the random forest, I found a few ways I believe the modeling process could be improved:

- Larger dataset
  - Better class balance
  - Pretrained BERT model
- 

# Recommendations and Next Steps

Using the vectorizer built for the Random Forest model, I was able to create lists of the most similar words in the vector, to keywords like “Apple,” “iPhone,” and “iPad,” for example.

1. **Apple Store** - One of the most positive words associated with Apple is 'Store' / 'Applestore.' People love the Apple Store, it was revolutionary when first introduced. How can you market the store and highlight how great it is, and get people to come in?
2. **Personal Device** - Some common Positive iPhone words are 'handy' and 'schedules.' People love how great the iPhone is as a personal device for day-to-day tasks. Create marketing to highlight businessmen using the iPhone for scheduling meetings, calendar appointments, etc. Or families scheduling playdates for children and soccer games etc.
3. **Sales** - A recurring negative word is sales. Prices are high and there aren't enough discounts and sales for iPhones, iPads, etc.

# Contact Information

Cassidy Exum

[cassidy.j.exum@gmail.com](mailto:cassidy.j.exum@gmail.com)

