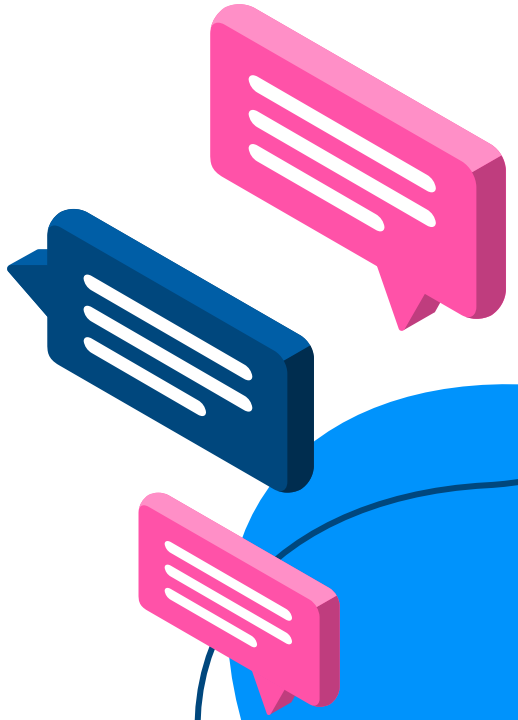
Twitter Language Processing

Alex Zieky, Ben Spilsbury, David Shin



Our Business Goals

A. Effectively Predict Tweet Sentiment

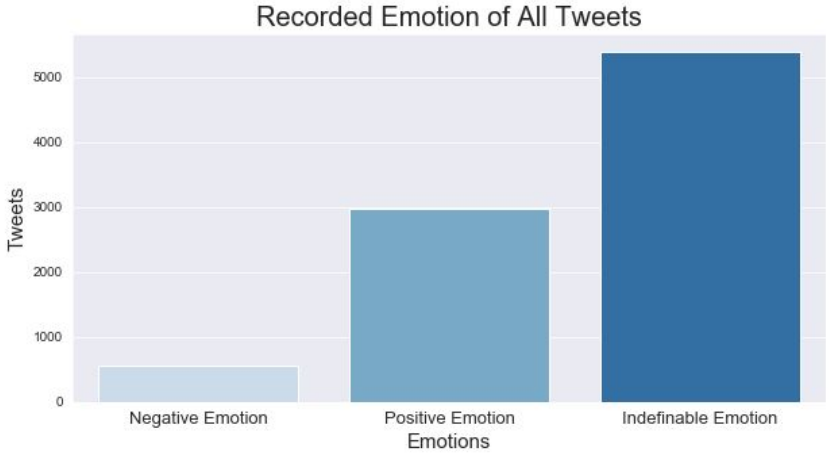
B. Analyze Tweets Directed at Apple & Google

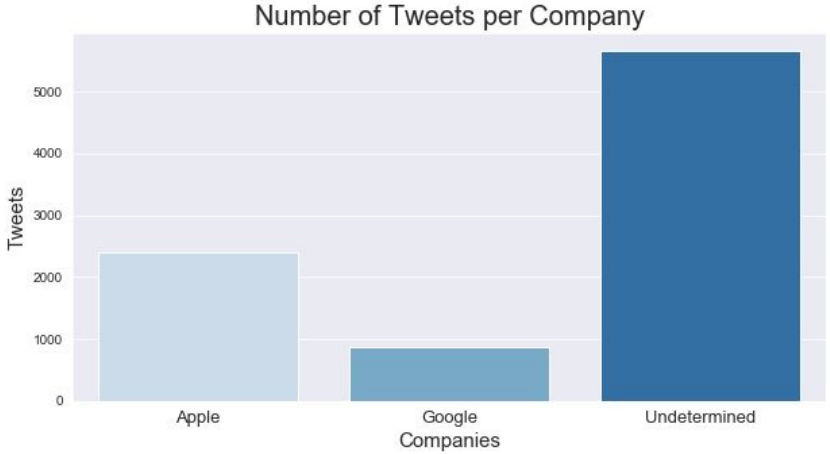
Data

- Dataset from CrowdFlower via dataworld
- 9,000 Tweets
- Target Variable: Emotion toward Brand

General EDA

- Target Variable EDA
- Companies
- Takeaways





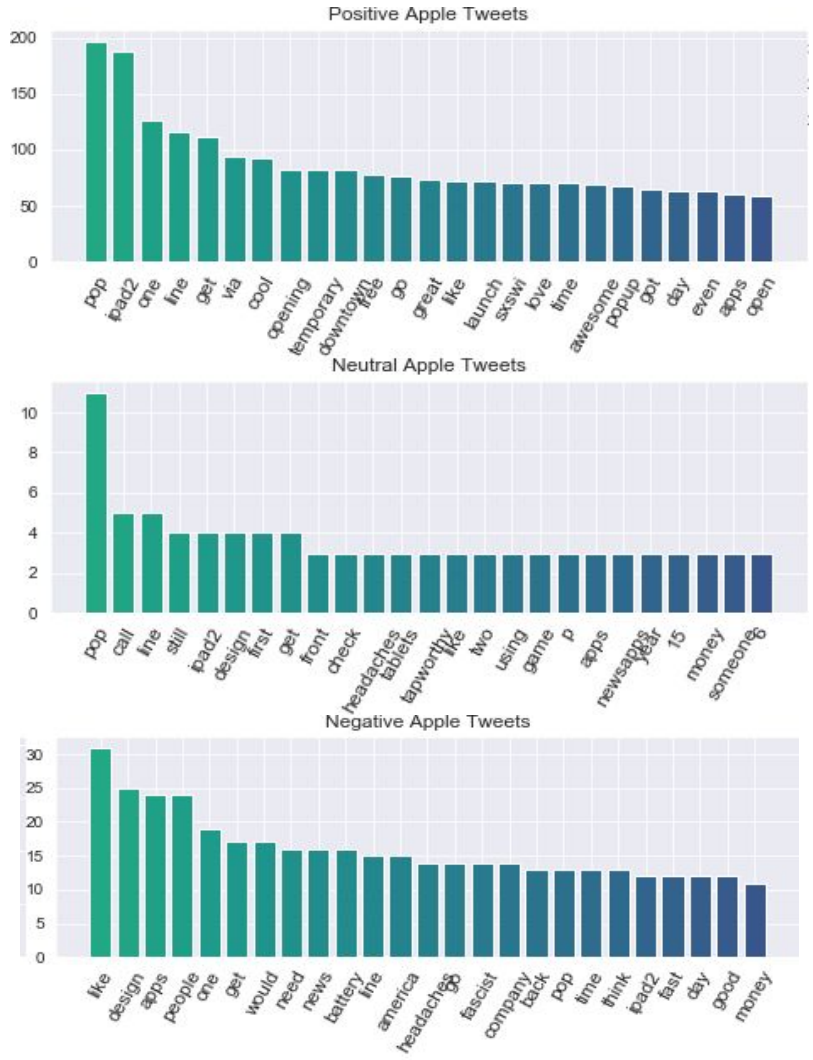
Word Clouds

Positive Words

Negative Words

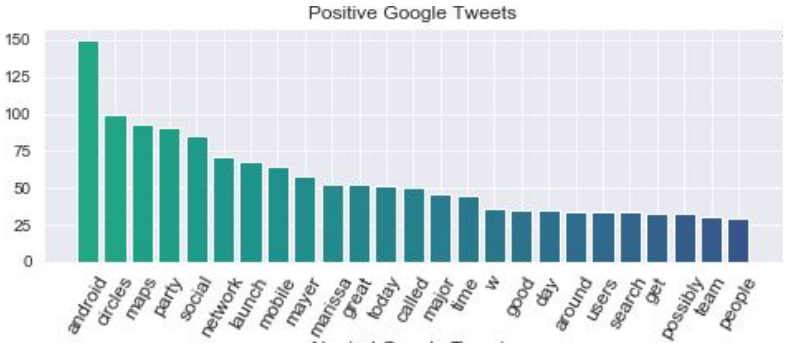
Apple EDA

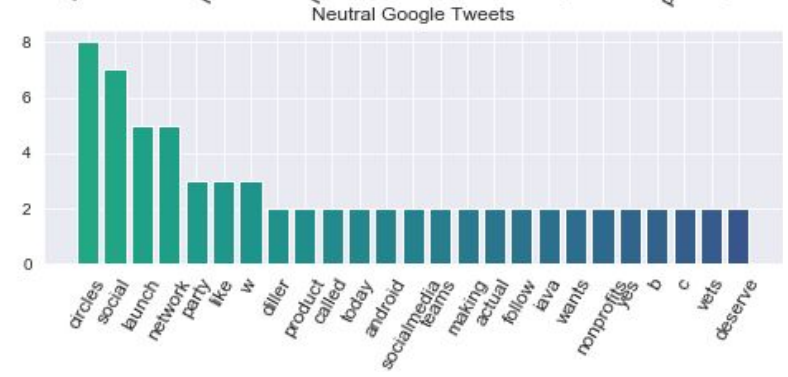
- 2,340 Tweets
- Apple Topics
- Positive Words
- Negative Words

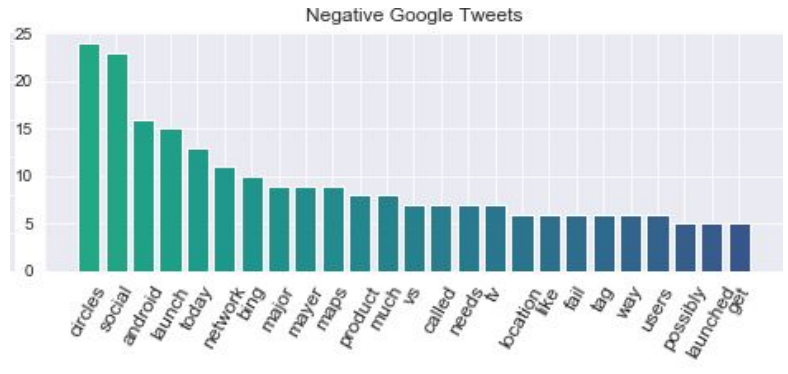


Google EDA

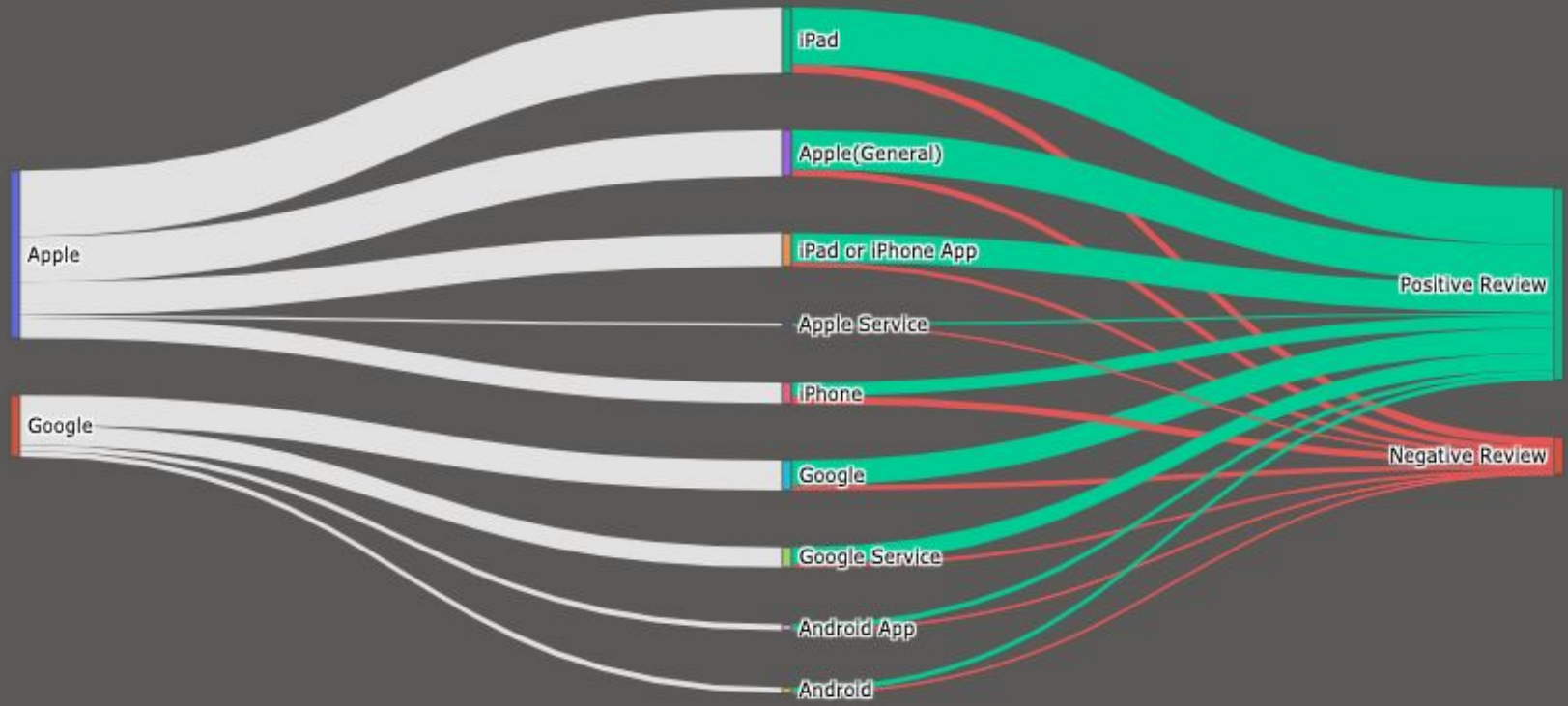
- 854 Tweets
- Google Topics
- Positive Words
- Negative Words







Apple VS Google: Twitter Reviews



Modeling Methodology

Clean/ Lemmatize

Train/Test Split

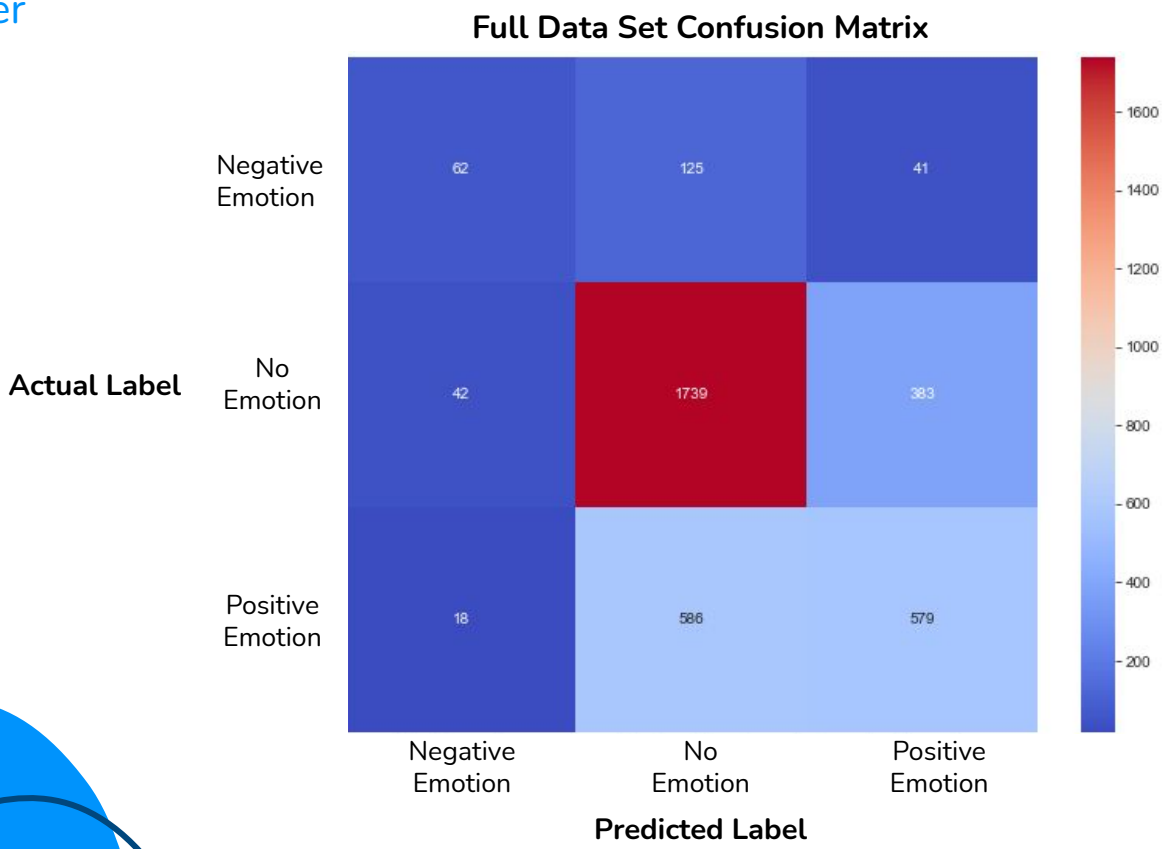
Model

Evaluate

Full Data Set Model

Support Vector Classifier

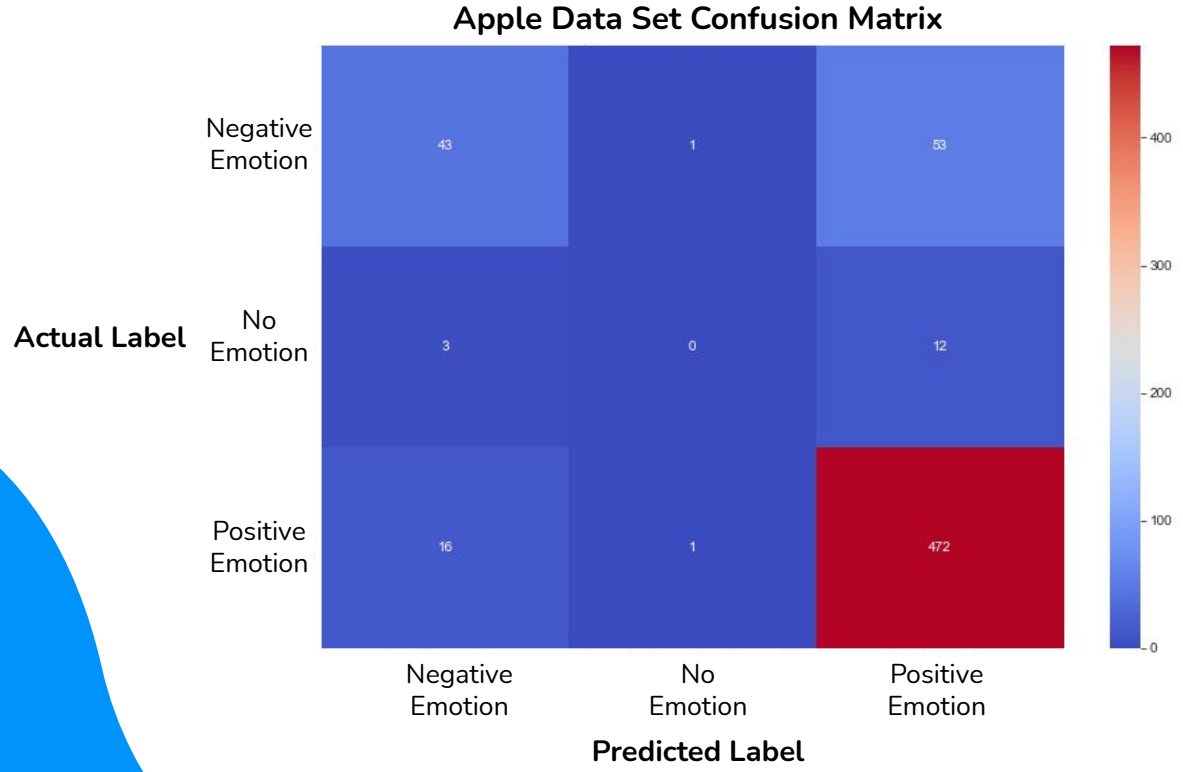
- Accuracy: 66.5%
- Multi Class
- Key Takeaways



Apple Model

Support Vector Classifier

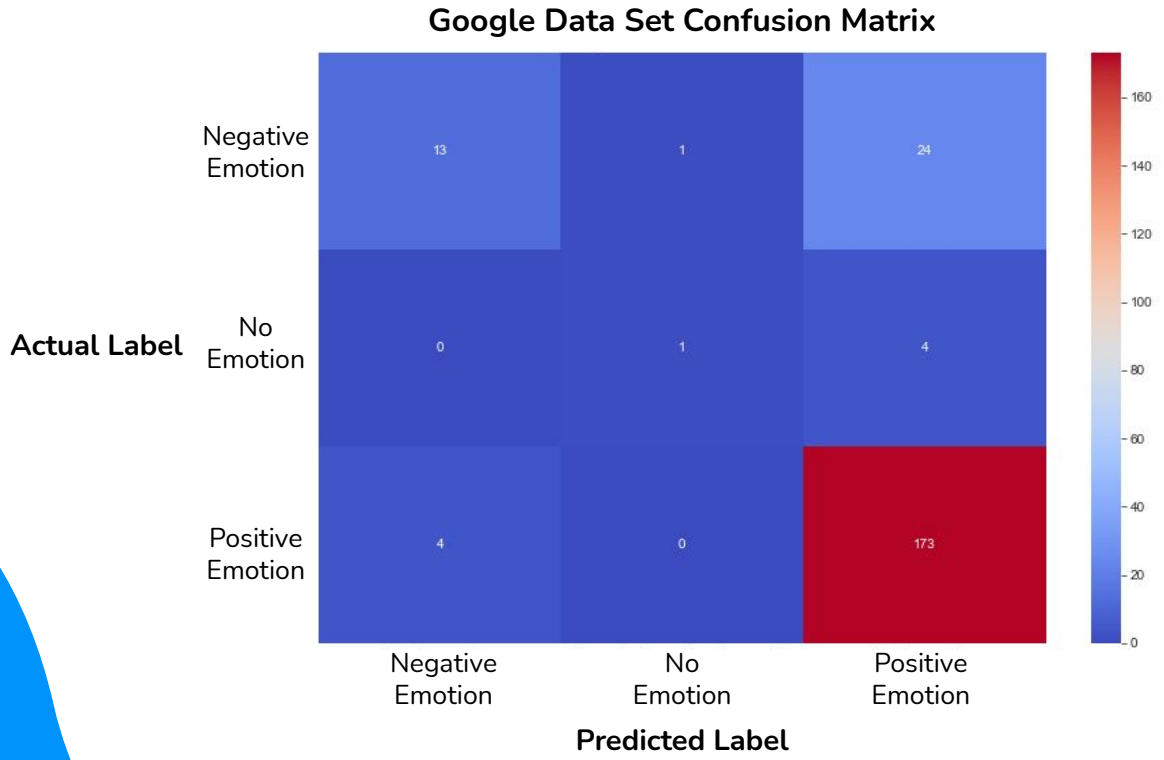
- Accuracy: 85.6%
- Key Takeaways



Google Model

Support Vector Classifier

- Accuracy: 85%
- Key Takeaways




CornFusionMatrix
@CornFusionMatrx
How could I forget my cell phone today for #SXSW what a fail glad I have my iPad #lifelinetotheworld
12:00 PM · Jun 1, 2020

Exceptions

Conclusion and Next Steps

- Model Performance
- Key Takeaways
- Future Implementation

Thank You!