Twitter Sentiment Analysis

Vijaya Gunukula

Goal

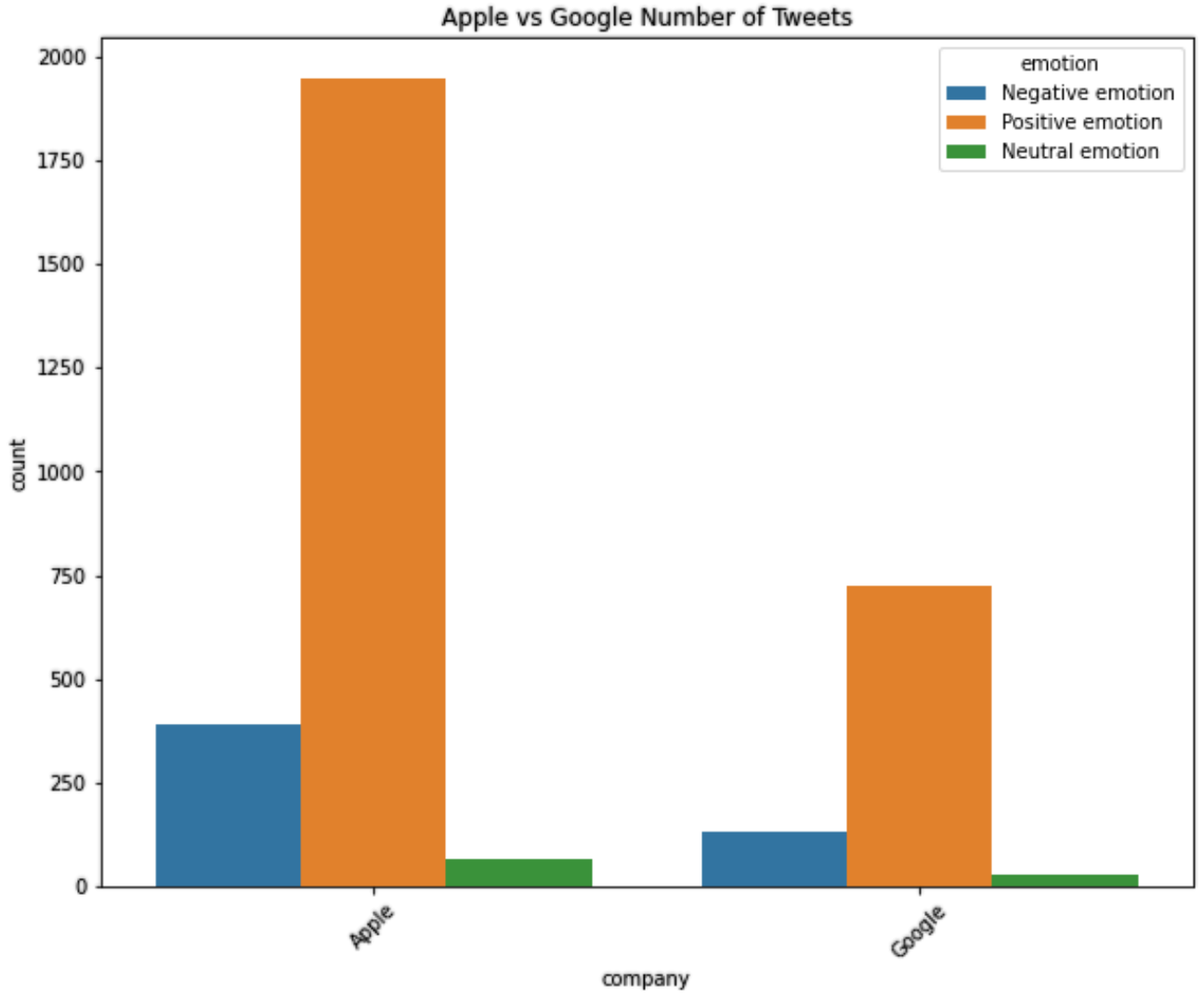
- The goal of this project is to perform supervised sentiment analysis on twitter tweets about google and apple products using Natural Language Processing and the model can rate the sentiment of the tweet based on its content.

Data

- Obtained data from CrowdFlower
- Source: https://www.crowdflower.com/data-for-everyone/
- This dataset has tweets of human raters rated the sentiment in over 9000 Tweets as Positive, Negative or Neutral.

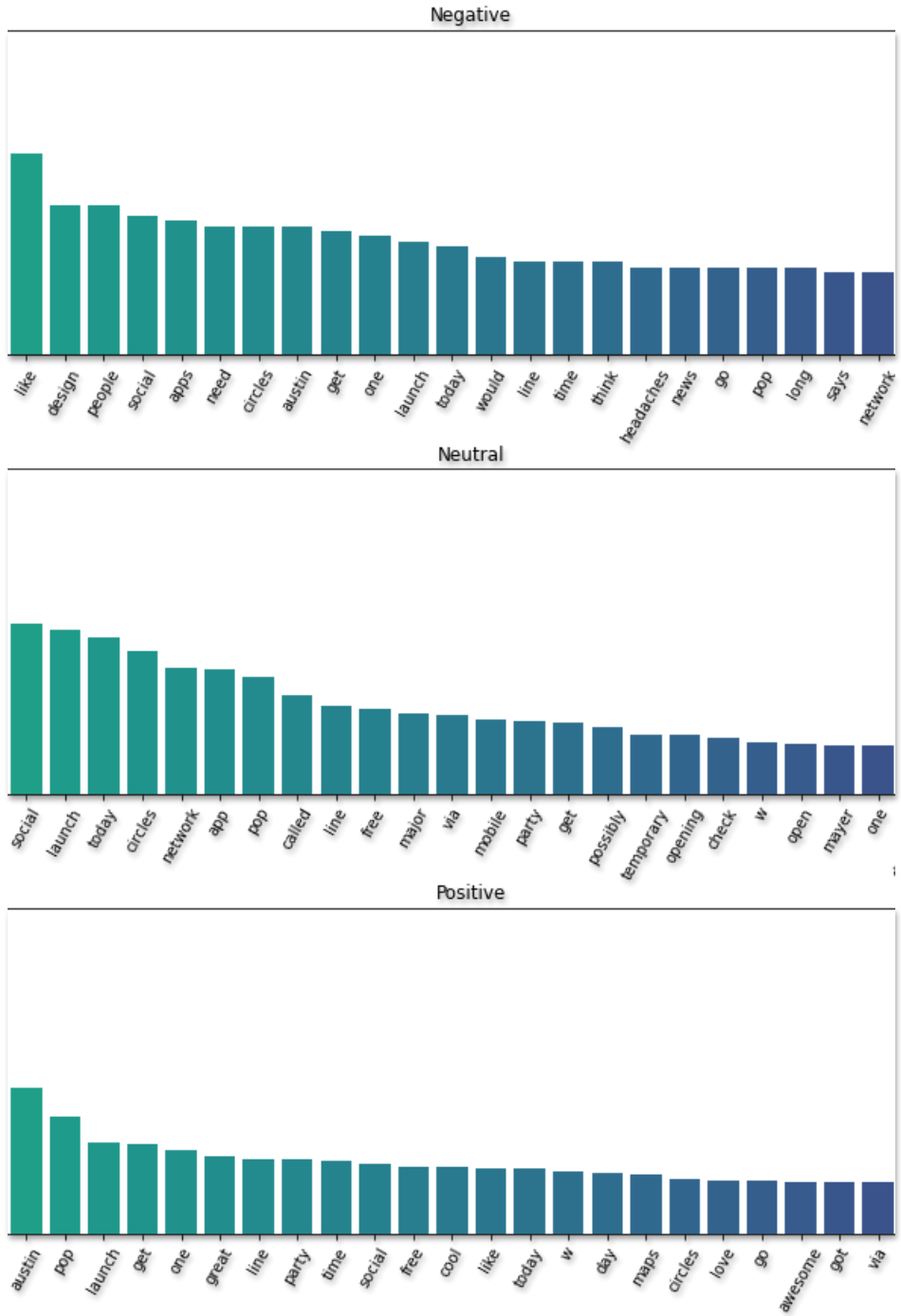
Number of tweets toward Apple and Google

- More positive tweets on both companies that negative and neutral tweets
- Apple has more tweets than Google



Top 30 words in tweets

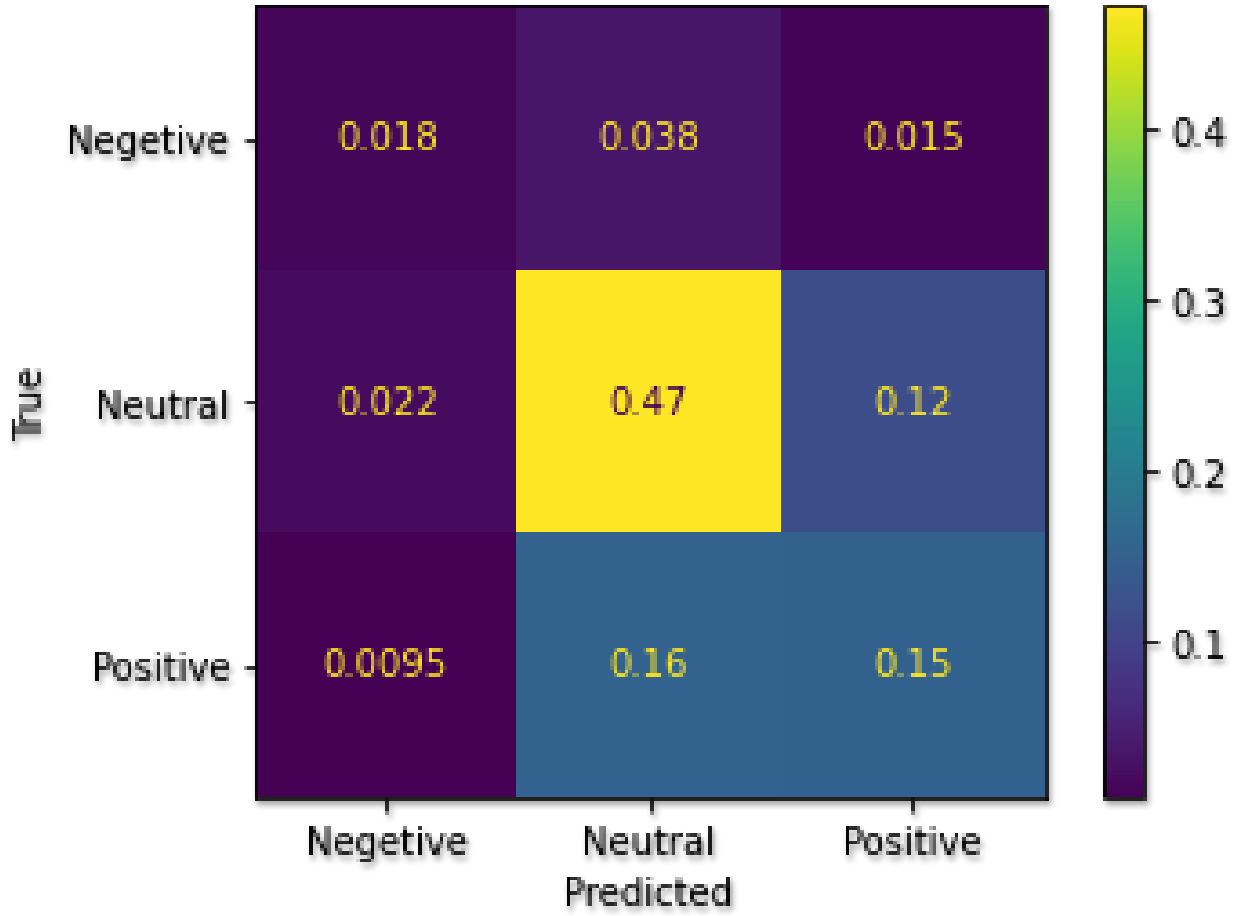
The frequencies of the top 30 words in Positive, Negative and Neutral tweets



Model

Naive Bayes model as Accuracy –64%

Confusion of Matrix of multiclass of the Naïve Bayes.



Deep NLP Model

Accuracy

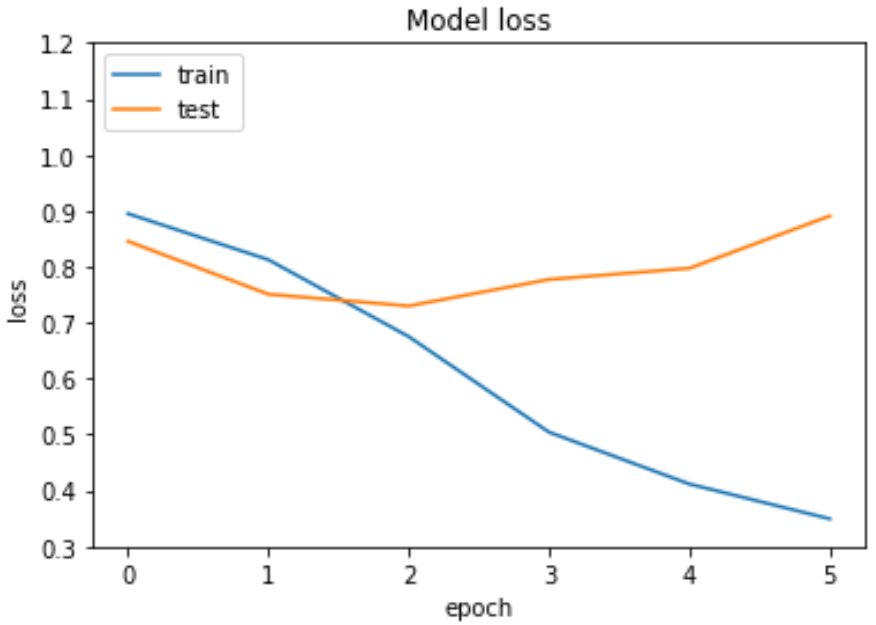
Train set : 87%

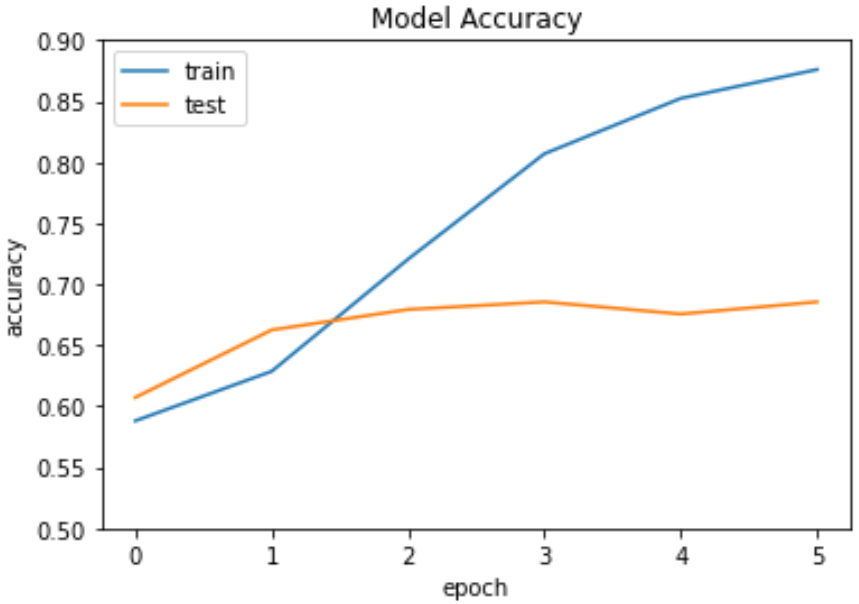
Test set : 66%

Loss Function

Train set : 0.36

Test set : 0.97





Conclusion

Models performs better on training set than test set which leads to overfitting, to avoid this problem difference should be minimum.

Future Work

- Adding more embedding layers and hyper parameters in the neural network model to avoid overfitting.

Thank You

- Vijaya Gunukula
- Email: vijayared@yahoo.com