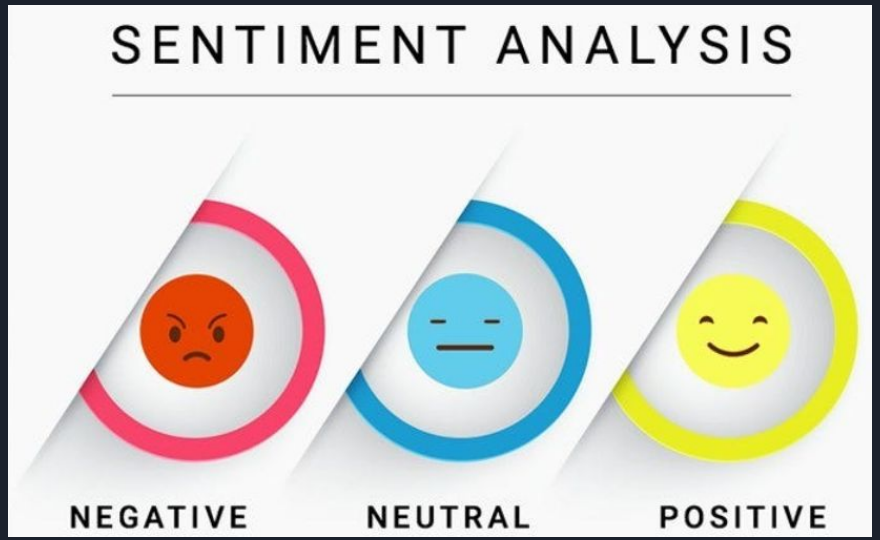
Twitter sentiment analysis using NLP

By: Greg Fatouras

Business Problem

- How can companies leverage social media to get a better understanding of how people view their products?
- How can companies determine customer satisfaction regarding products or services?
- What is the product sentiment before and after a marketing campaign?

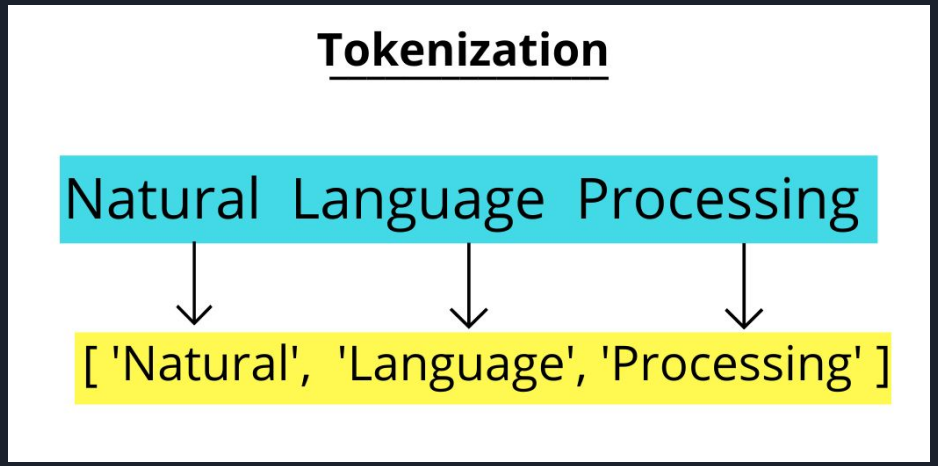


Dataset

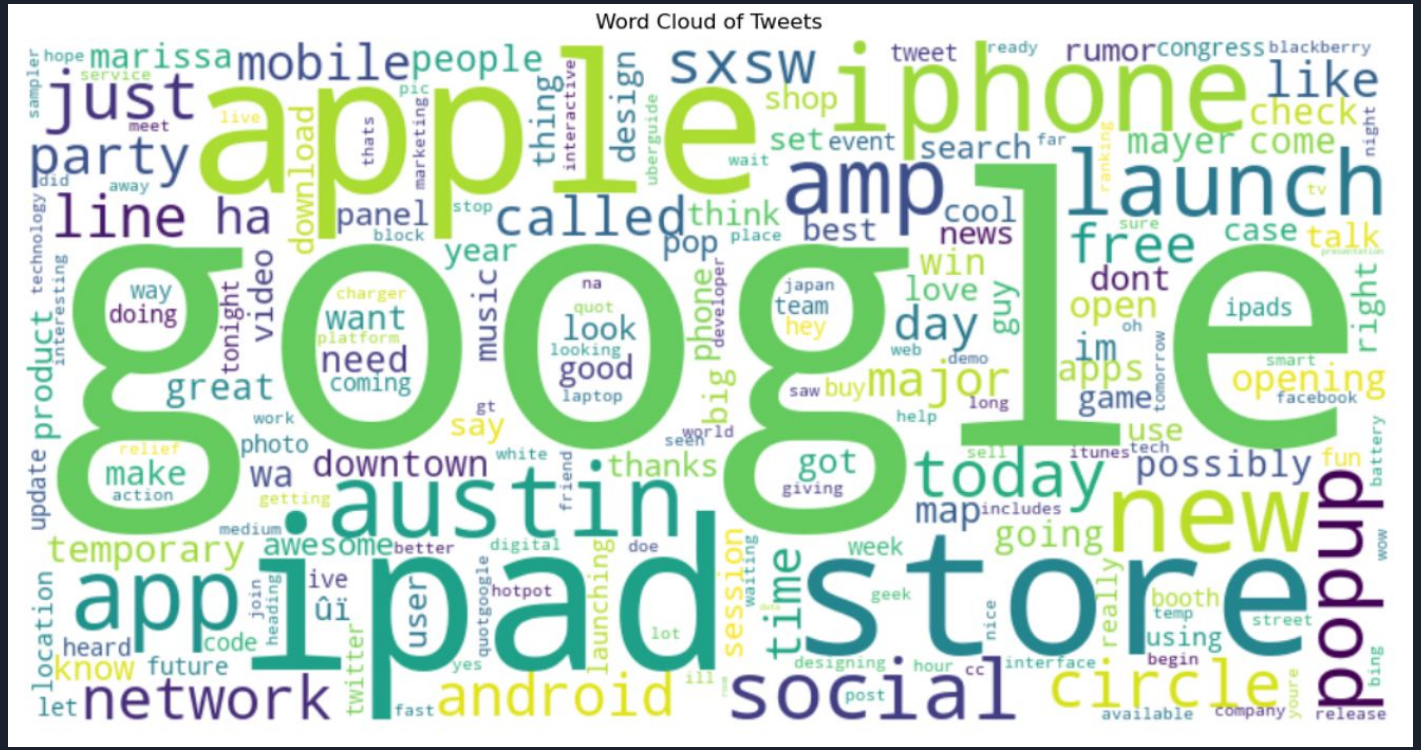
- 8,909 tweets
- Which company/product are the tweets directed at
- The sentiment of the tweet, (Positive, Neutral, Negative)

Methods of tackling problem

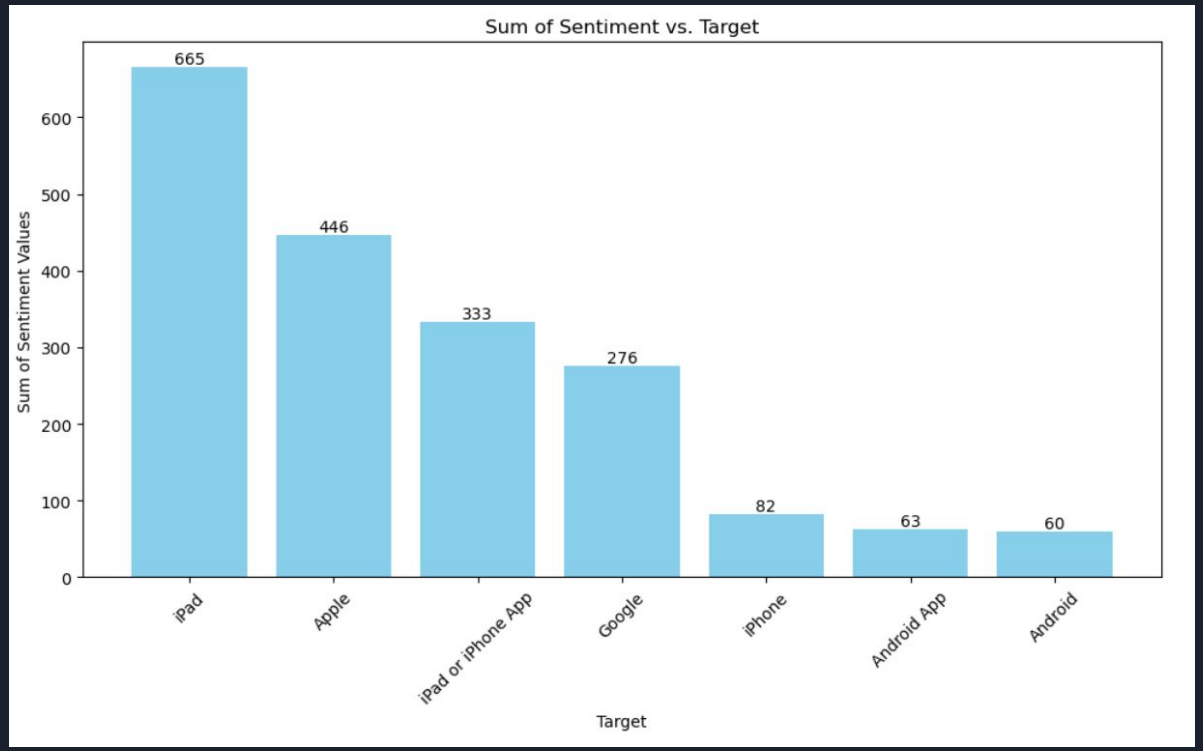
- Split each tweet into individual words
- Tie each tweet to a sentiment (Positive, negative, netural)
- Balance the sentiment categories for optimal classification performance
- Feed information to a machine learning model



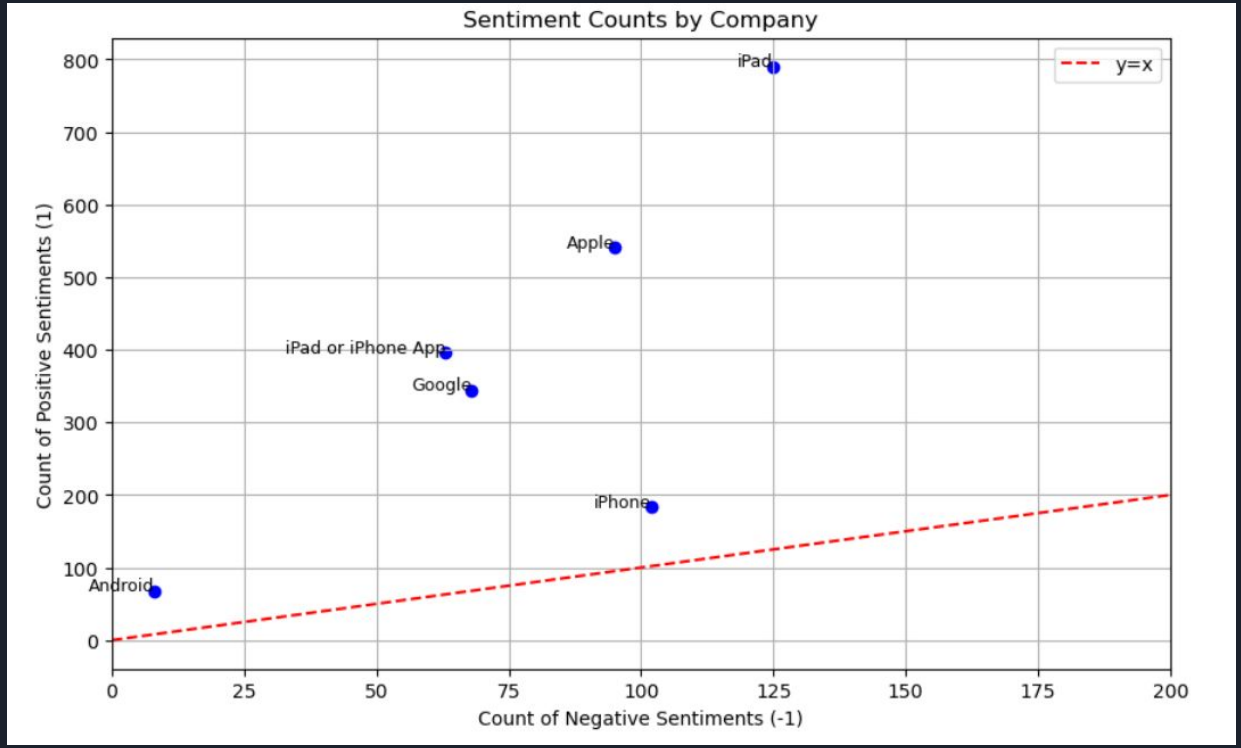
Word Frequency



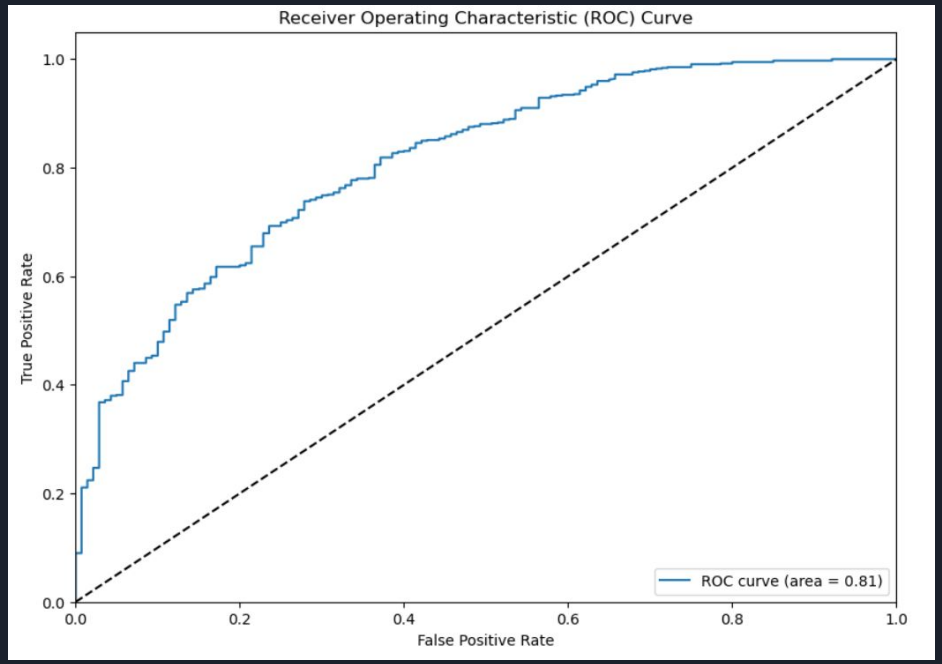
Average Sentiment vs. Company / Product

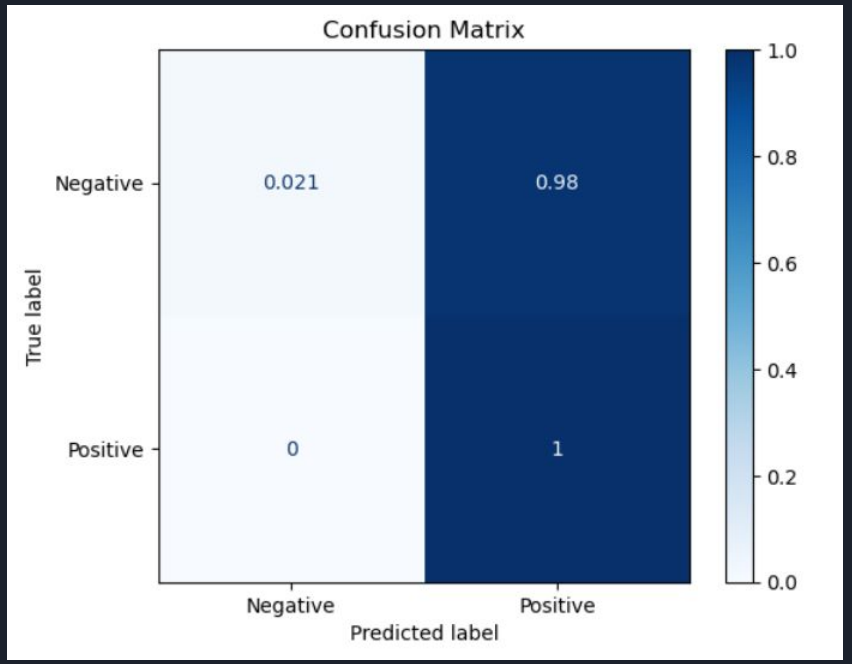


Average Sentiment vs. Company / Product



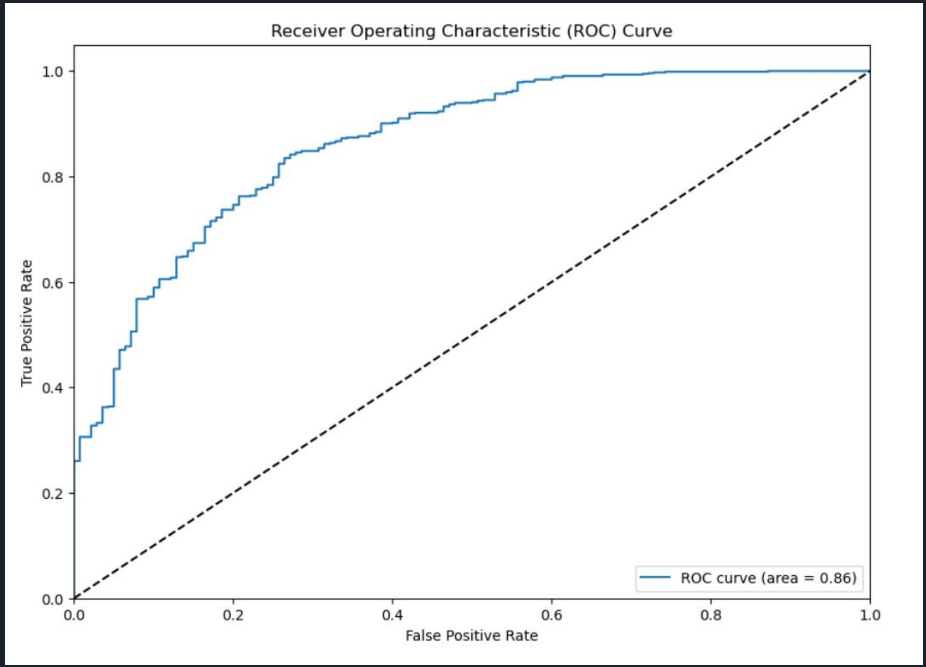
Baseline Model (MultinomialNB)

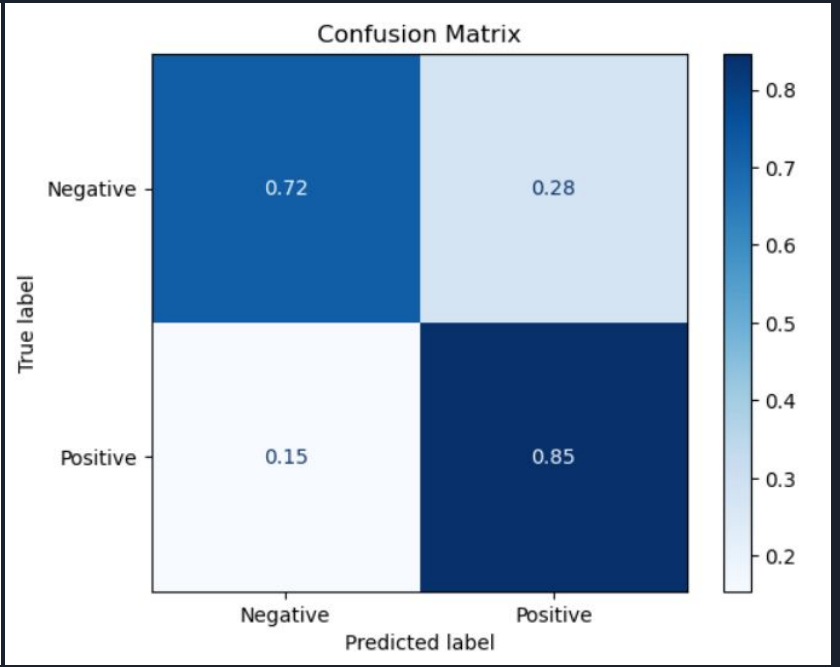




Accuracy: 0.85*

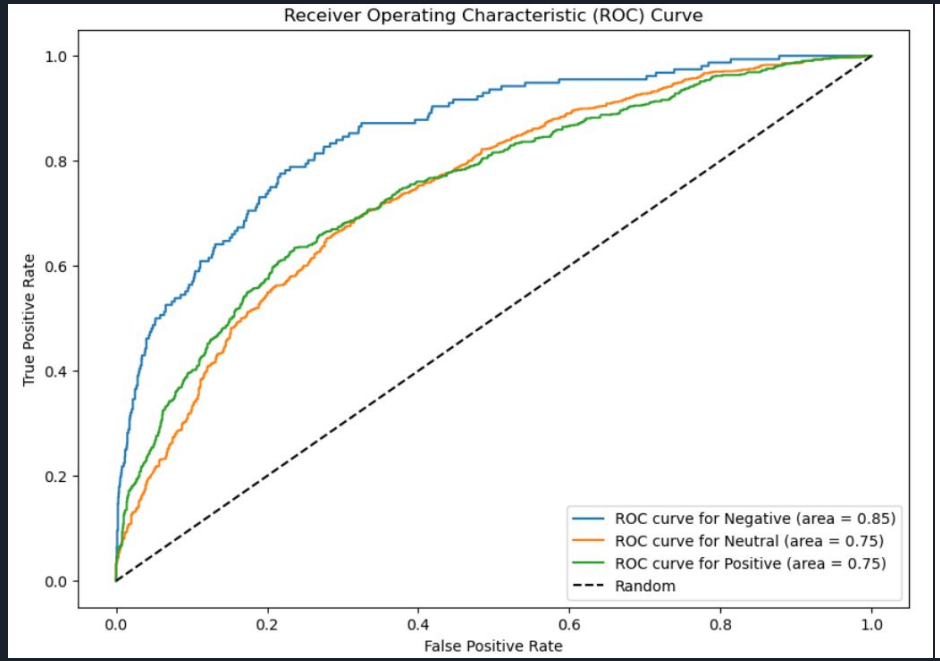
Tuned Binary Model (MultinomialNB)

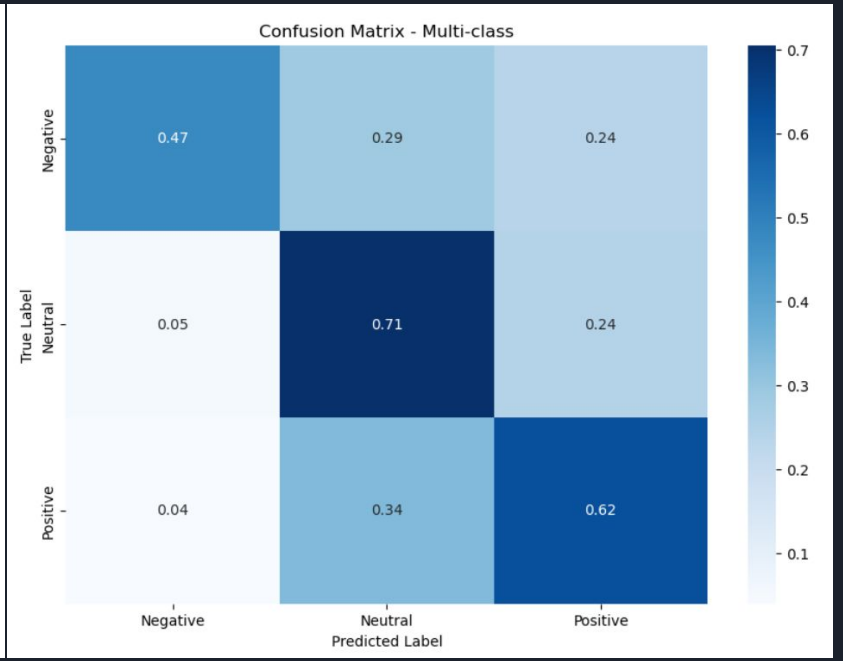




Accuracy: 0.83

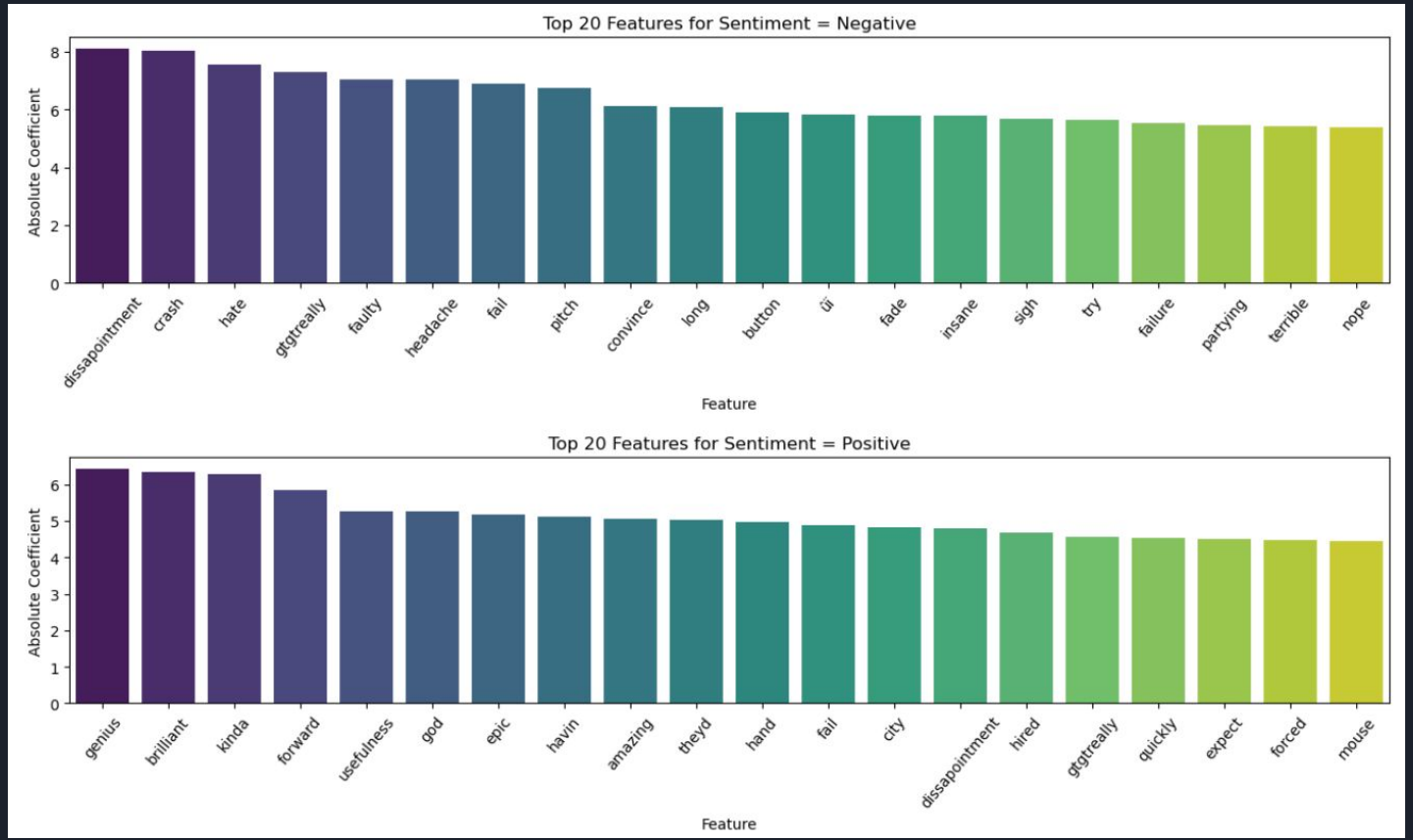
Tuned Multi-class Model (MultinomialNB)





Accuracy: 0.66

Significant Features



Recommendations

- Collect data using keywords that relate comments to the business.
- Ensure that there is an equal mix of positive and negative sentiment within the dataset.
- Use the MultinomialNB model to determine positive or negative sentiment
    - Use the model to determine the success of marketing events
    - Compare sentiment before and after competitor releases

Conclusion

- The Multi-Class MultinomialNB model struggled with differentiating positive and neutral tweets
- Neutral tweets are not very useful for determining sentiment towards a product.
- Binary MultinomialNB accurately differentiates positive and negative sentiment

Thanks You

Greg Fatouras

- https://github.com/Gfatouras
- https://www.linkedin.com/in/gfatouras/
- fatourasg@gmail.com