NLP: Sentiment Analysis

Analyzing Product Sentiments from Tweets

Sung Bae

The Importance of Communication

"Our beliefs and perceptions of reality, and the choices we make, are largely conditioned on how others see and evaluate the world"

- Bing Liu , University of Illinois, Chicago

Twitter

145 million daily users

22% of Americans are on Twitter

500 million tweets sent each day

65.8% of US companies use Twitter for marketing

80% of Twitter users have mentioned a brand in a tweet

77% of Twitter users feel more positive when their tweet has been replied to.

Goals

Create a model that can

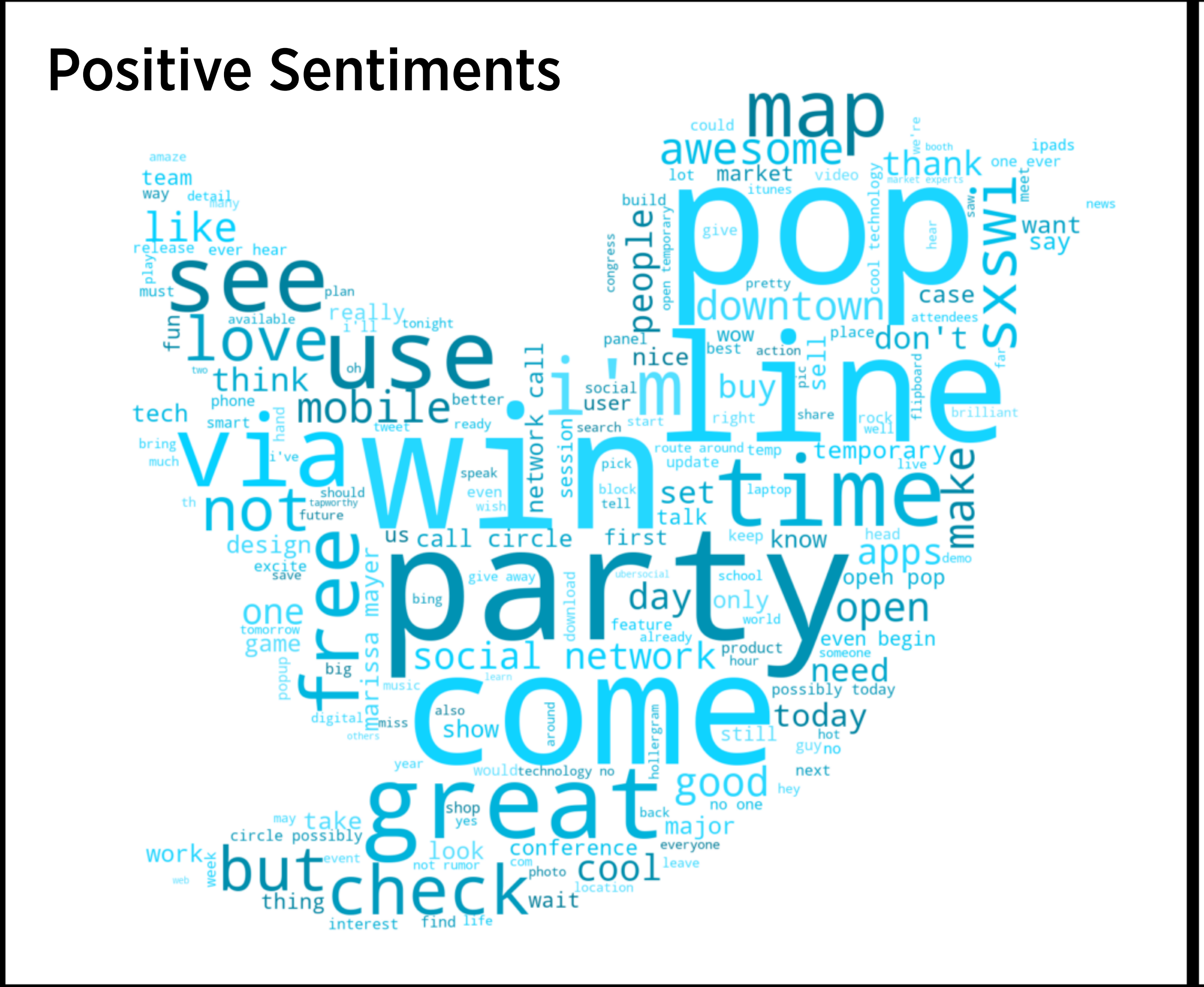
[1] Flag the company when there is a negative tweet about their specific product

[2] Correctly identify which product

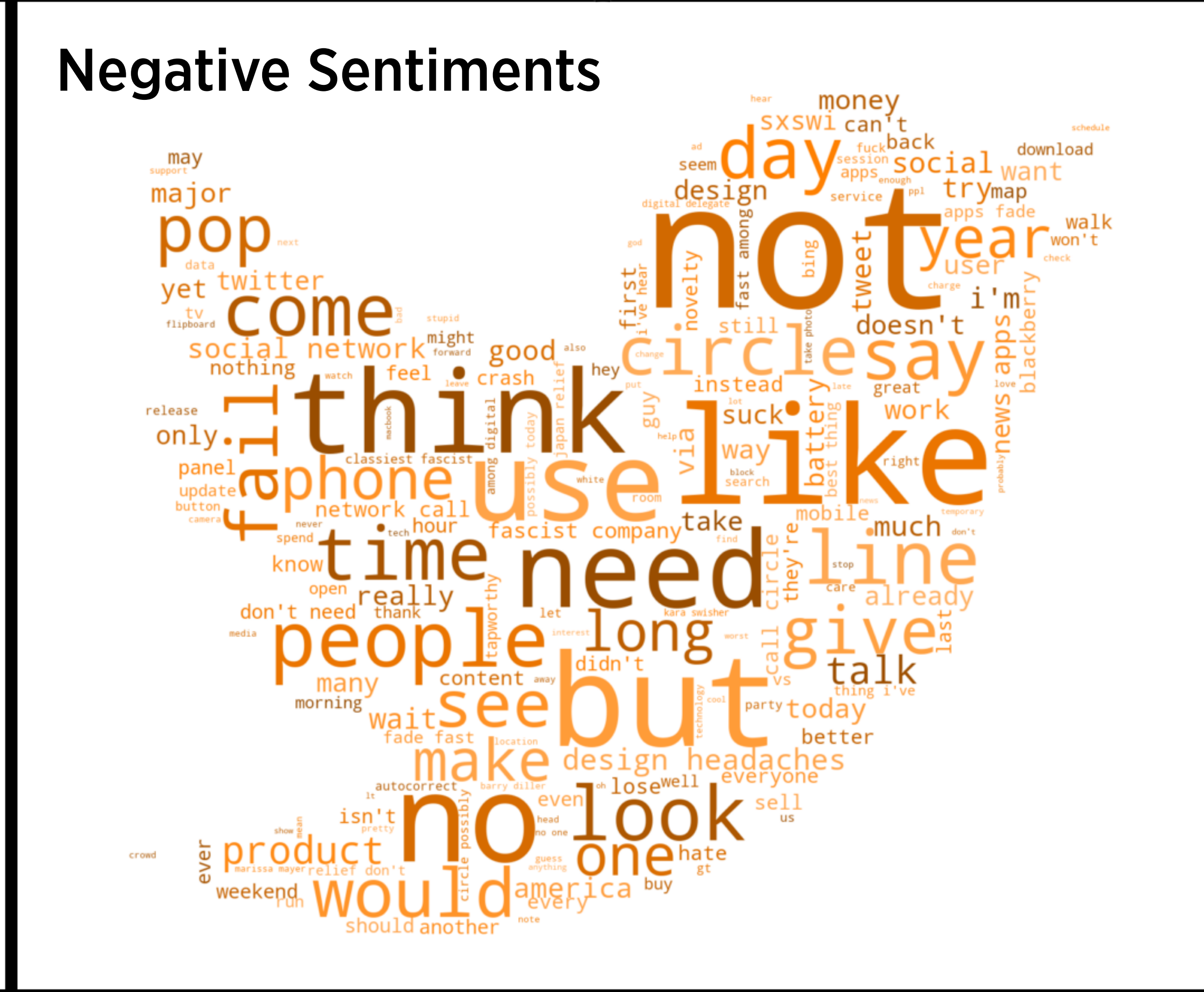
[3] Correctly classify tweets into negative, neutral, and positive sentiments

EDA: Positive and Negative Sentiments

Word Cloud: Positive Sentiment



Word Cloud: Negative Sentiment



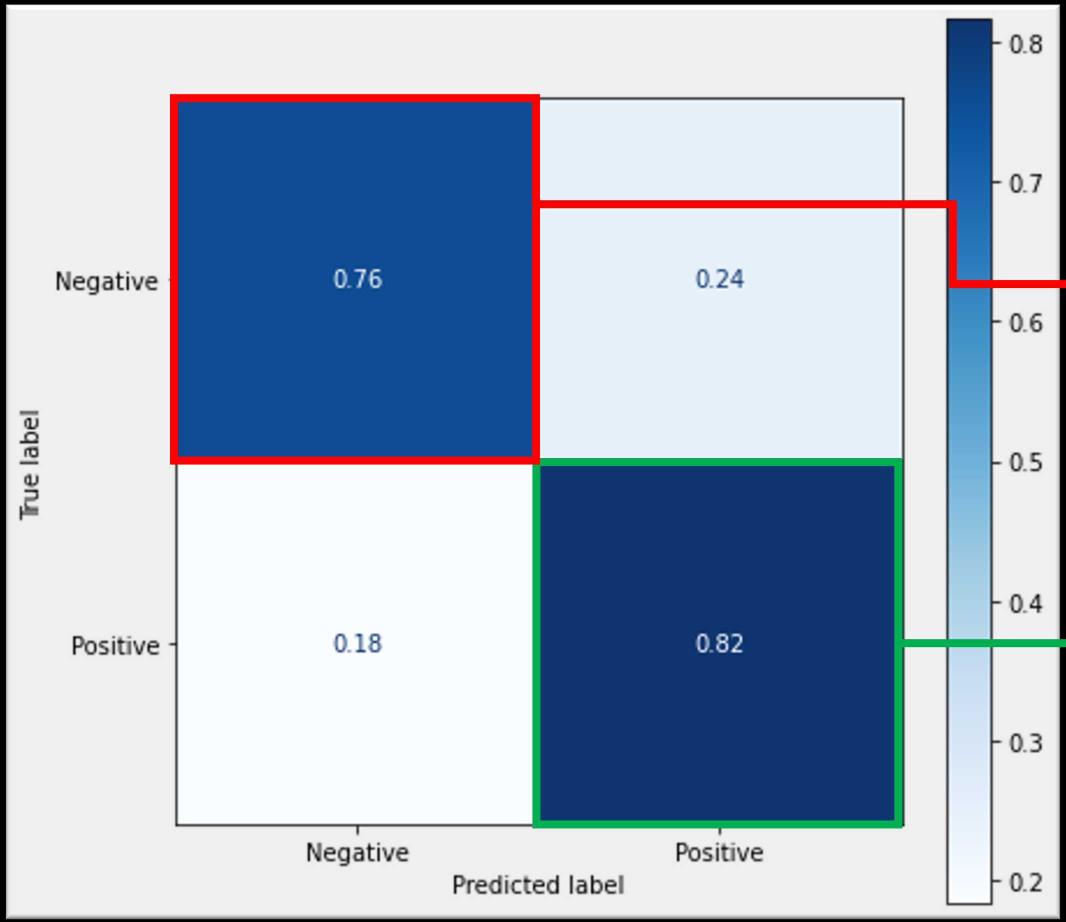
Binary Models Result

label	negative_recall	positive_recall	test_accuracy	average_time
LR_spacy_we	0.737226	0.828000	0.813980	0.018581
NB_nltk_tfidf_smote	0.708029	0.830667	0.811725	0.000474
LR_nltk_tfidf_smote	0.642336	0.893333	0.854566	0.000508
SVM_nltk_tfidf_smote	0.605839	0.925333	0.875986	0.000522
NB_spacy_tfidf_smote	0.540146	0.597333	0.588501	0.011910

Best Negative Recall:
[LR_spacy_we]
Logistic Regression - Word embedding

Best Computing Time:
[NB_nltk_tfidf_smote]
Naïve-Bayes' – TF-IDF with SMOTE

Model Results: Binary (Positive vs. Negative)

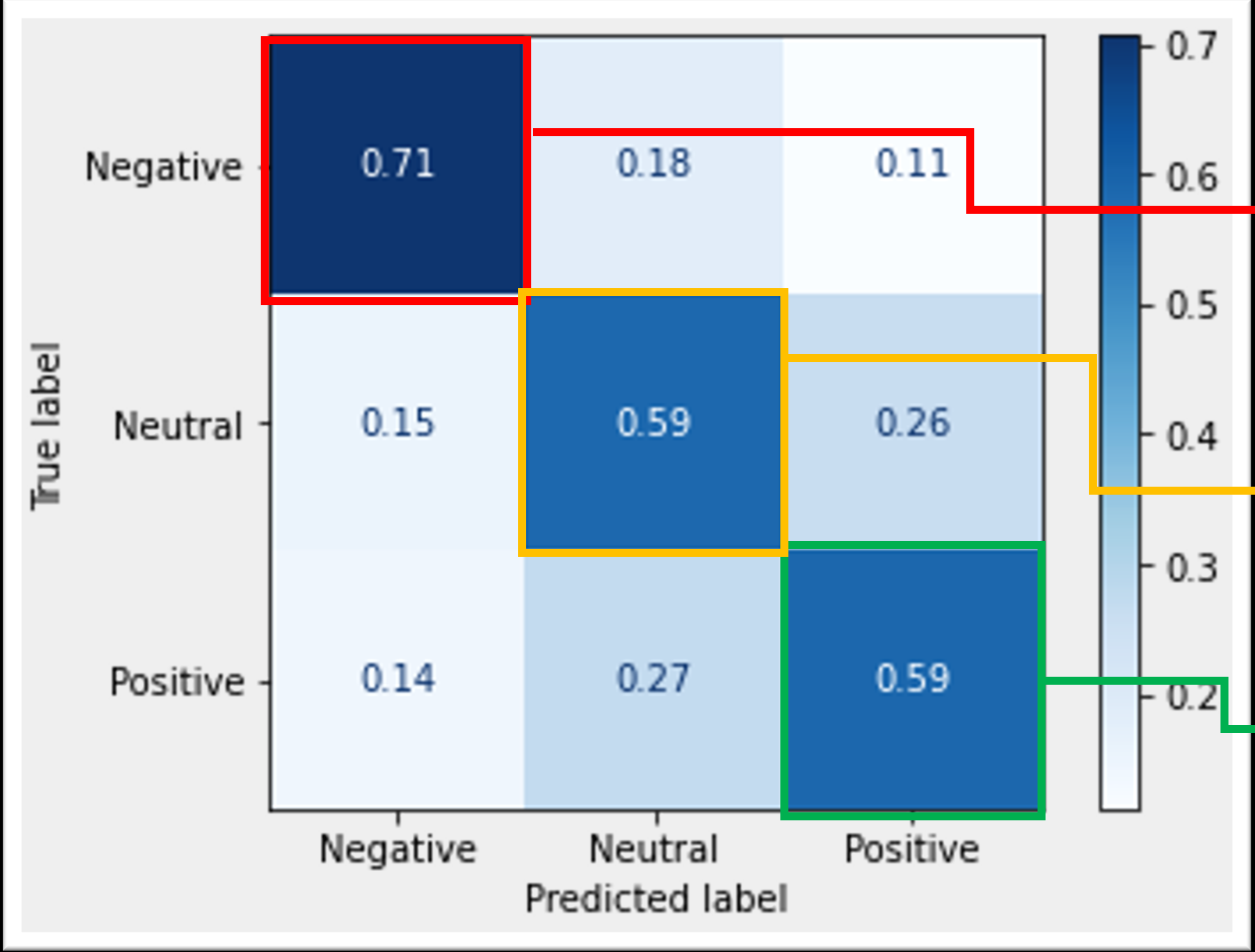


Overall Accuracy = 81 %

76% of negative sentiment tweets were correctly identified

82% of positive sentiment tweets were correctly identified

Model Results: Multiclass

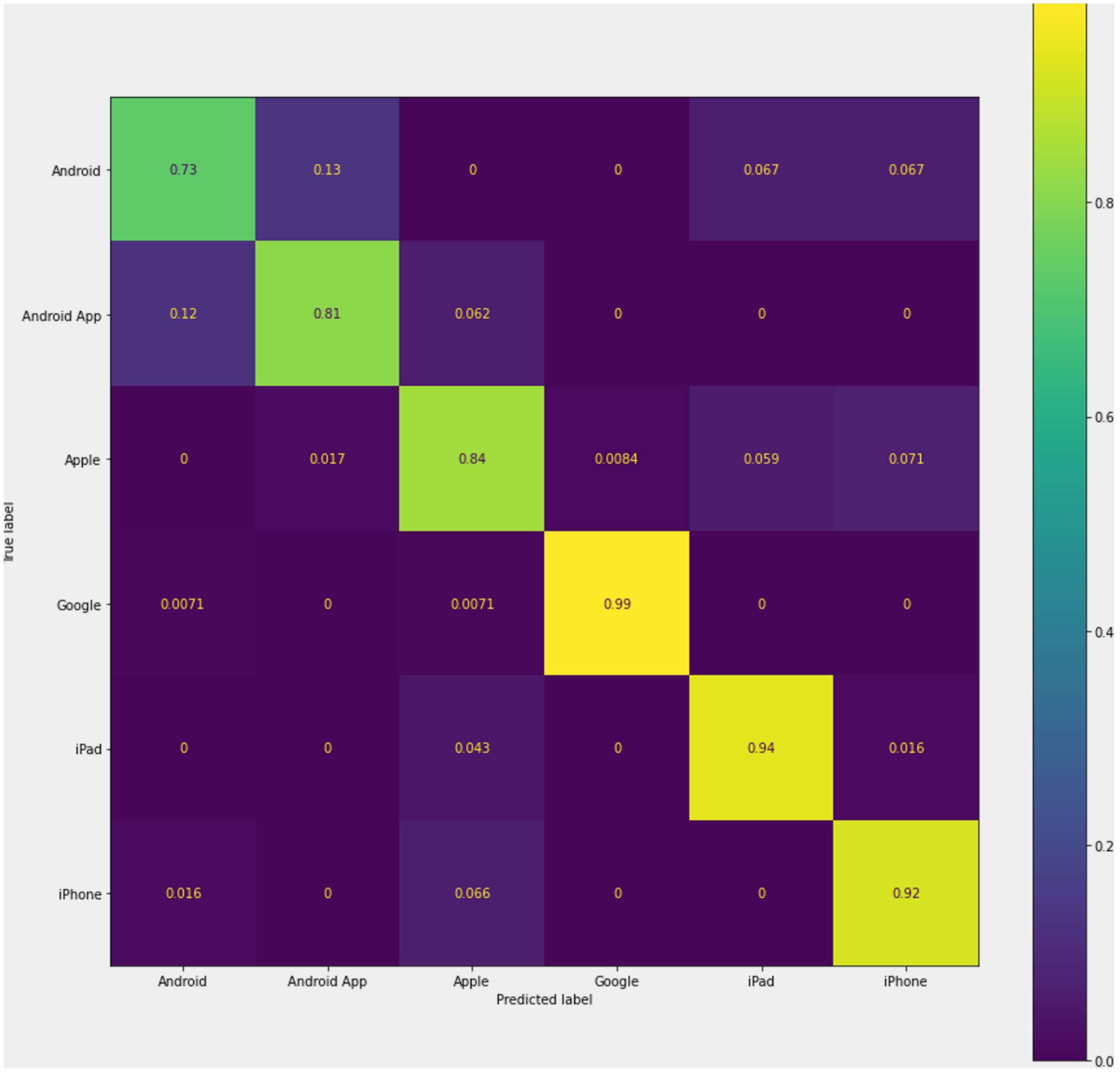


Overall Accuracy = 60 %

71% of negative sentiment tweets were correctly identified

59% of neutral sentiment tweets were correctly identified

59% of positive sentiment tweets were correctly identified



Model Results: Product Predictor

Overall Accuracy = 91 %

Comment:
Products under same brand are confused by the model

Example.
Android and Android App

Conclusion

- Our final models can
  - Binary Sentiment
    - 81% accuracy with 76% negative sentiment recall
  - Multiclass Sentiment
    - 60% accuracy with 71% negative sentiment recall
  - Product Predictor
    - 91% of accuracy

Recommendations 1 – Binary Sentiments

- Multilcass not only complicates the model itself, but also it dilutes negative and positive sentiments together, therefore, we recommend using binary sentiment classification model

Recommendations 2 – Models

1. Best Negative Sentiment Identifier:

logistic regression using word-embedding model [74%]

Pros:

- Best Negative Recall (74%)
- Good Overall Accuracy (81%)

Cons:

- Takes longer time (40x longer)

---

2. Best Computing Time Sentiment Identifier:

Naïve Bayes’ using TF-IDF model with SMOTE

Pros:

- Best Computing Time
- Good Negative Recall (70.8%)

Cons:

- Uses SMOTE – Synthetically generated data

*Further studies need to be done for consistency

Future Direction

[1] Further data acquisition:

- Due to imbalance in data, our model suffered. Acquiring similar datasets with more balanced classes would increase our model's performance and reliability.

[2] Real time Twitter analysis

- Have this model refined and put in a production so that it can monitor tweets real time and analyze and negative flags about any particular product.

[3] Research various costs between models

- Different models have different strengths and weaknesses, and we believe researching different costs including computing and labor costs using different models would allow us to choose a better model

[4] Try different NLP techniques

- There are so many different NLP techniques and they are growing due to works of so many different research entities. It would be very valuable to try different approach including skip-gram and even deep learning NLP for this project.

Thank you for listening

Appendix