Project: Sentiment Analysis Classification

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions and Next Steps

Business Problem

- One E-commerce company would like to understand customer sentiments on Apple and Google products.
- They made a request to perform sentiment analysis on the tweet text on Twitter
- We work on this project to build classification predictive models

Business Problem

Business insights to investigate:

- Predict if a twitter text indicates Positive/Negative/Neutral emotion

Data

Source:

- Dataset from Data Word

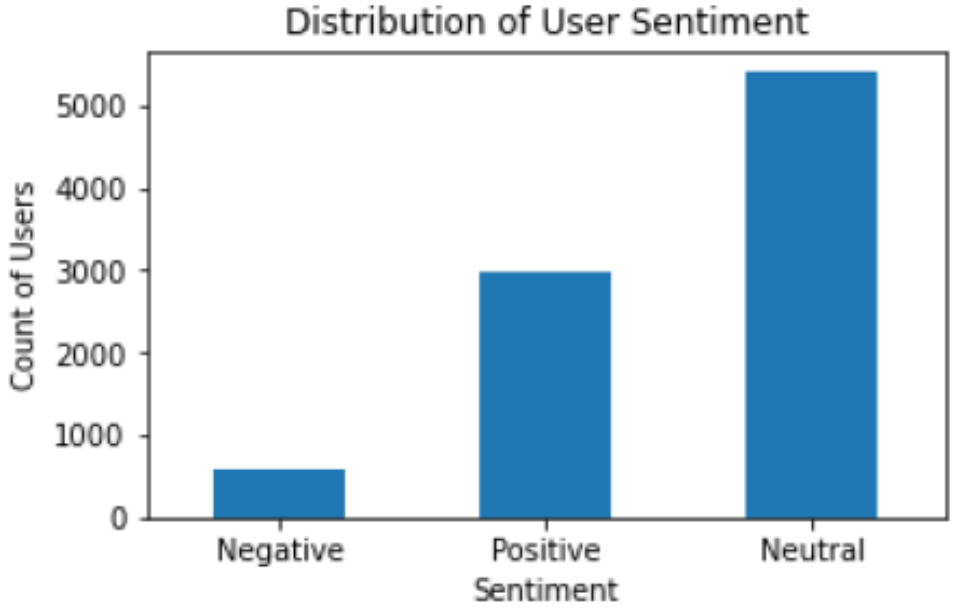
Information:

- Twitter text about the product
- Emotion Labels for the comment

Methods

- Classification Approach for NLP applications
- Train Classification Model for Sentiment prediction
- Evaluate Model Performance

Results

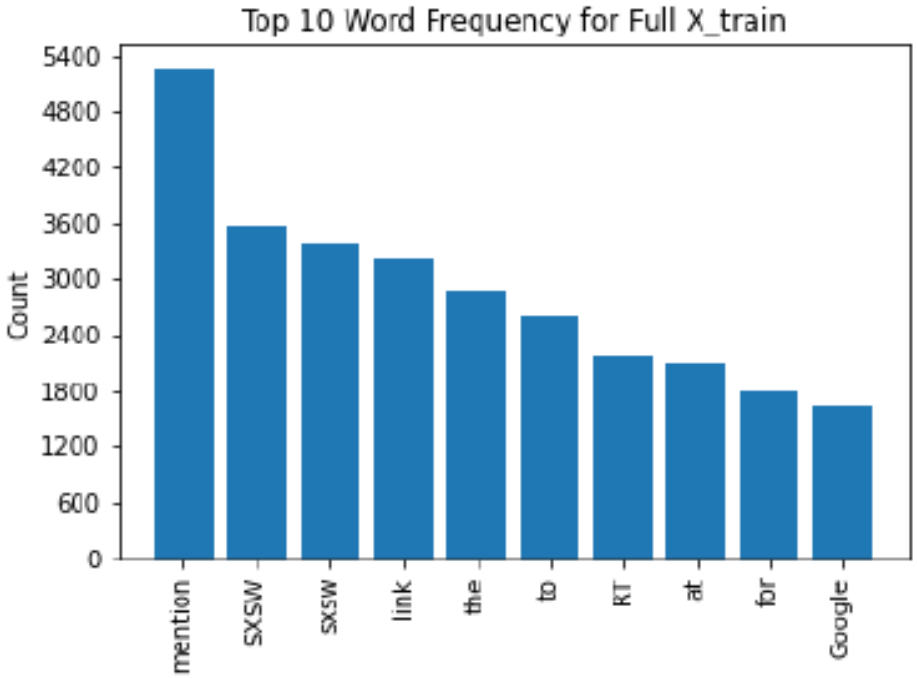


9000 Entries of Data

User Sentiment Labels:
Neutral: 5389
Positive : 2978
Negative: 570

Results

Exploratory Data Analysis



Top 10 Frequent Words:

- Mention, SXSW, sxsw, link, the, to RT, at, for, Google

Results – Modeling

Model : Classification Models

Feature Engineering : Tfidf

- Base model :
  - Naïve Bayes Model, 10 Max_Features
- Second Model :
  - Naïve Bayes, Stemmed Tokens + Stop Words Removal + 100 Max_Features
- Final Model:
  - Naïve Bayes, Stemmed Tokens + Stop Words Removal + 250 max_Features

- Model improvement with accuracy

Results

	Model	Accuracy
0	Baseline	0.608774
1	100 Features, Stemming	0.641450
2	250 Features, Stemming	0.650850

Model Performance

- Accuracy : 65%

Conclusions

Classification models for User Sentiment Classification

- Model prediction for multiclass classification outcome

Recommendations

- We had developed classification models for prediction. Further research and development are needed to improve model accuracy for practical application.

Next Steps

- Generality of Model Performance
  - Limitation : Simple Model, Low Accuracy
- Feature Engineering for Model Improvement
- Advanced Model
- More data

Thank You!

Email: aliang@gmail.com
GitHub: @username