Be Smart Twitter

Zihan Zhou
Nicholas Montoya
Ziyang Yang

Questions

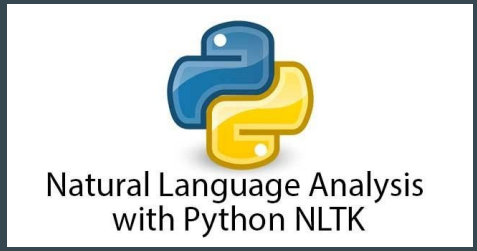
- Are tweets with more ‘positive’ and ‘negative’ words more closely to people’s real moods?
- How does people’s mood change according to the time period of a day?
- How do people use the Emojis (like smile) express their real sentiments? Does the Emoji support their expressions or not?

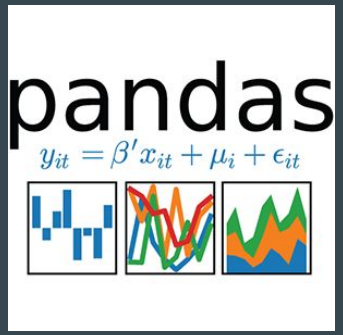
Data Preparation Work

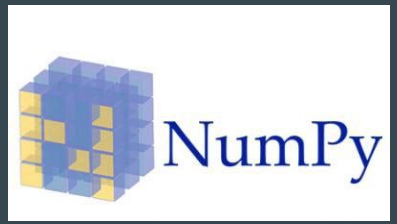
- Removing usernames
- Removing miscellaneous characters
- Tokenizing tweets
  - Vectorizing tweets
- Removing duplicate tweets from various data sources
- Separate specific times into ranges

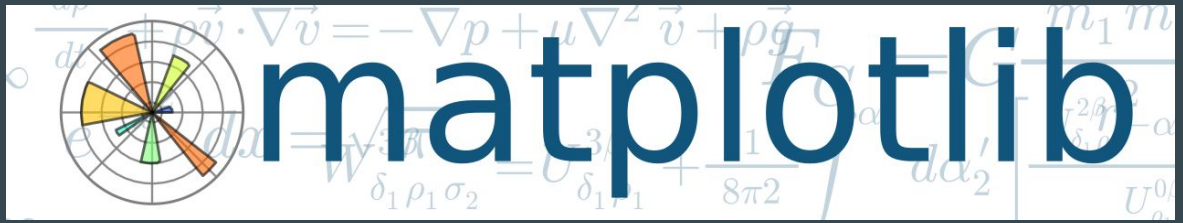
Tools used

- Natural Learning Tool Kit (NLTK)
- NumPy
- Pandas
- matplotlib



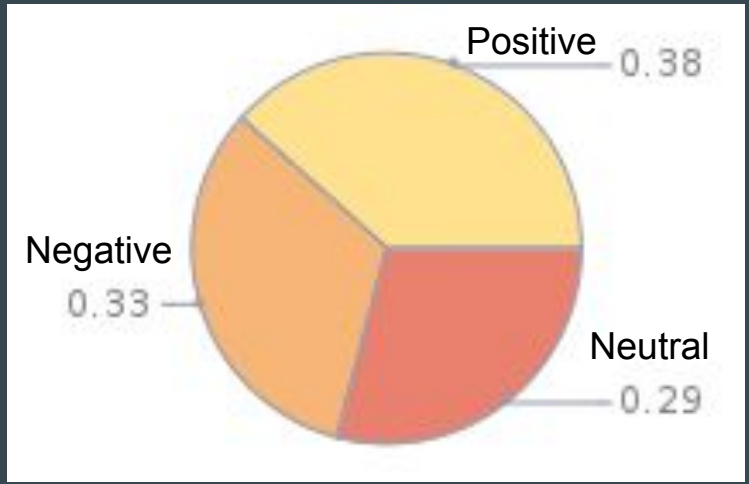






Classification & Clustering

- Made SVM to classify tweets
- Classified tweets as positive, negative, or neutral along with subjective vs. objective using Vader (NLTK)
- K-Means clustering with various ‘K’



Knowledge Gained

- Positive sentiment intensity peaks in the morning
- Negative sentiment is relatively constant throughout the day
- Neutral sentiment peaks during daytime hours but intensity of others rises
- Sentiment around shopping changes from morning to night
- Emojis/Emoticons are easily misinterpreted due to irony and sarcasm
- It's more possible that people use negative words in negative sentiment than using the positive words in positive sentiment

Application

- Target advertisements between 7:30-9pm
- Release top headlines during standard work hours
- Improve language processing libraries
- Collate and classify reviews (like | dislike | neutral)
- Analyze and predict political elections