# Sentiment Analysis

IMDB Action Movies Reviews

Kevin Okiah

4/4/2019

# **Sentiment Analysis**

**Problem Statement**: Can Sentiment Analysis be used to Isolate Positive and Negative Movie reviews?

**Sentiment Analysis** seeks to gauge opinions, feelings, Attitudes, or Emotions about a given subject from written or spoken language. Is it Positive or Negative.





Positive

### **Results**

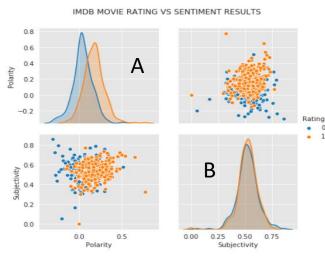


Fig. 1

#### Word in Positive Sentiment



#### Words in Negative Sentiment



#### **Data Set and Tools Used**

- Source: IMDB Action movies Reviews
- Number of Movies: ~300 movies
- Type: 1 Positive Review and 1 Negative Review per Movie
- Data Collection Method: Web-Scrapping with Beautiful Soup, LXML,
- Data Preprocessing Included: Tokenization, Changed to lower, contractions expanded, Stop words and punctuations removed, lemmatization, and vectorization using TF-IDF for Clustering.
- **Sentiment:** TextBlob used for Polarity [-1,1] and Subjectivity [0,1] Sentiment calculations

## **Results Interpretation and Conclusion**

- P Sentiment Polarity: Positive Reviews (1) have a more positive polarity distribution compared to Negative reviews(0). Movie Rating is correlated to Sentiment Conveyed by reviewer. Positive Rating is directly related to Sentiment. See Fig. 1(A)
- Sentiment Subjectivity: There is no difference in distribution of Subjectivity of Positive(1) and Negative(0) reviews. Fig. 1(B)
- Positive Sentiment Words include 'Good', 'best', 'great' etc.
- Negative Sentiment Words include 'bad', 'worst', 'really' etc
- Conclusion: Sentiment Analysis can be used to Isolate
  Positive and Negative reviews