Sentiment Analysis using NLP

GROUP MEMBERS :

P.V. CHANDRA SHEKAR

G. SAMHITHA

N. DINESH

K. RAHUL

K. SASHANK

Problem Statement :

Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material, and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations.

Key Applications :

- Social media monitoring
- Customer support management
- Analyzing customer feedback
- Market research and competitive research

Sentiment Analysis with NLP on Twitter Data

Md. Rakibul Hasan1, Maisha Maliha2, M. Arifuzzaman3

Every social networking sites like Facebook, Twitter, Instagram etc become one of the key sources of information. It is found that by extracting and analyzing data from social networking sites, a business entity can be benefited in their product marketing. Twitter is one of the most popular sites where people used to express their feelings and reviews for a particular product. In our work, we use twitter data to analyze public views towards a product. Firstly, we have developed a natural language processing (NLP) based pre-processed data framework to filter tweets. Secondly, we incorporate Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) model concept to analyze sentiment. This is an initiative to use BoW and TFIDF are used together to precisely classify positive and negative tweets. We have found that by exploiting TF-IDF vectorizer, the accuracy of sentiment analysis can be substantially improved and simulation results show the efficiency of our proposed system. We achieved 85.25% accuracy in sentiment analysis using NLP technique.

Natural Language Processing, Sentiment Analysis and Clinical Analytics
Adil Rajput Assistant Professor

Recent advances in Big Data has prompted health care practitioners to utilize the data available on social media to discern sentiment and emotions' expression. Health Informatics and Clinical Analytics depend heavily on information gathered from diverse sources. Traditionally, a healthcare practitioner will ask a patient to fill out a questionnaire that will form the basis of diagnosing the medical condition. However, medical practitioners have access to many sources of data including the patients' writings on various media. Natural Language Processing (NLP) allows researchers to gather such data and analyze it to glean the underlying meaning of such writings. The field of sentiment analysis – applied to many other domains – depend heavily on techniques utilized by NLP. This work will look into various prevalent theories underlying the NLP field and how they can be leveraged to gather users' sentiments on social media. Such sentiments can be culled over a period of time thus minimizing the errors introduced by data input and other stressors. Furthermore, we look at some applications of sentiment analysis and application of NLP to mental health. The reader will also learn about the NLTK toolkit that implements various NLP theories and how they can make the data scavenging process a lot easier.

Multilingual Twitter Sentiment Classification: The Role of Human Annotators
Igor Mozetic, Miha Grcar, Jasmina Smailovic

What are the limits of automated Twitter sentiment classification? We analyze a large set of manually labeled tweets in different languages, use them as training data, and construct automated classification models. It turns out that the quality of classification models depends much more on the quality and size of training data than on the type of the model trained. Experimental results indicate that there is no statistically significant difference between the performance of the top classification models. We quantify the quality of training data by applying various annotator agreement measures, and identify the weakest points of different datasets. We show that the model performance approaches the inter-annotator agreement when the size of the training set is sufficiently large. However, it is crucial to regularly monitor the self- and inter-annotator agreements since this improves the training datasets and consequently the model performance. Finally, we show that there is strong evidence that humans perceive the sentiment classes (negative, neutral, and positive) as ordered.

Tentative Plan :

Before 2nd review :

     Data Set Collection

     Data Pre-processing

Before Final review :

     Implementation

     Improving Efficiency of the Model