Twitter Sentiment Analysis

Twitter Sentiment Analysis is analysis of text sentences and get the sentiment out of it. Nowadays this is very important because it is very effective way to predict the current trend going on in the society. People are more expressive on social media like twitter, than they are when giving feedbacks to certain service they used.

Using Sentiment Analysis, one can find the reviews of certain thing, whether positive or negative. Also, some predictions could be made, like Who is going to be the next President of the country.

For sentiment analysis, the basic idea is split the sentence into words and then compare those words with some dictionary words and find out whether those are positive or negative. More positive words mean more positivity and more negative means more negativity. We can also score the sentence based on the positive/negative words we found. Extreme negative sentences will be taken as anger of society and positive will be taken as success of something.

The complexity involved here is interpreting the positive sentences said with negative words. For ex: Apple is not that bad company. Here no positive word is there but the sentence is positive. Second challenge is identifying sentences said in sarcastic way.

Right now, the complexity is not considered, and a simple approach is mentioned in this project. First the sentence is broken into words and then they are compared with bag of positive and negative words. Score is given based on the positive and negative words are found. The positive and negative words are used from Lexicon library [1-2].

After studying the sentences, some words, which I felt are negative and positive and would help in recognizing the sentiment are added into the library.

Accuracy is tested like this:

In terms of pos and neg, it can be mentioned as below (Also calculated in program):

Neg Pos

Neg 261 55

Pos 49 114

Accuracy= Total correctly classified sentences/Total number of sentences

=(261+114)/479

=78.28%

Accuracy can be increased, by using the model which can identify sarcasm and sentences said in opposite way.

References:

[1].Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." ; Proceedings of the ACM SIGKDD International Conference on Knowledge ; Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, ; Washington, USA.

[2].Bing Liu, Minqing Hu and Junsheng Cheng. "Opinion Observer: Analyzing ; and Comparing Opinions on the Web." Proceedings of the 14th ; International World Wide Web conference (WWW-2005), May 10-14, ; 2005, Chiba, Japan.