

Sentiment Analysis using Natural Language Processing and Machine learning

Objective

- Be able to automatically classify reviews of various services as positive, neutral or negative
- Achieve high accuracy, recall and F1 score

Motivation

- Current techniques are not very accurate
- Finding the perfect solution is impossible, so there is always scope for improvement in any existing solution

Sentiment Analysis Process Flow

Problems

- Many solutions exist, but nothing is directly applicable
- For the sub-problems, there are many algorithms, techniques and data available
- The challenge is to choose the solution that will have the maximum positive impact on the overall system

Observations

- The current system has very low accuracy, especially for negative reviews
- The proposed system on the right has to be implemented, refined and tested

Future Work

- Experimentation and research till the optimal amount of accuracy is achieved
- Extensive testing using a variety of real data sets

prepared are very sloppy and time consuming. If I do not want an ingredient to be added (walnut for example), there is a sense of panic. The young guy who sometimes takes order has an attitude. This is one place where the customers actually initiate a smile and greeting and the guy doesn't even look up or respond.

Features/
Opinion
extraction
using
Aboot
Algorithm

["sloppy", "attitude",
"panic", "great",
"expensive", "delicious",
...]

Data
cleansing,
filtering,
and
stemming
using nltk

<Filtered opinion words>

+ 0.85

The food was great

0.00

The service was okay

- 0.75

The sandwiches were sloppy

Sentence/
reviews
classifier

["sloppy"] => - 0.75
["expensive"] => - 0.6
["great"] => + 0.85
.....

Sentiment scoring using WordNet
3.0, Senti-WordNet 3.0 and graph
techniques