**Sentiment Analysis**

**Introduction**

Sentiment analysis is the process of using natural language processing, text analysis, and statistics to analyse customer sentiment. Sentiment analysis algorithm analyses customer tone based on the review text. We use  statistics, and text analysis to extract, and identify the sentiment of words into positive, negative, abusive or neutral categories. Sentiment analysis can be used to measure the impact of new product, campaigns and generally user sentiments across different range of products.

**How Sentiment Analysis is calculated?**

Based on user entered review text, we create each word as token. We compare tokens against the list of positive, negative, inverters, exclamatory and abusive words. First we check number of tokens which are under positive, negative and inverters. We derive sentiment factor based on the below logic.

**Positive**: Check number of positive tokens. If there are 5 positive tokens and each token individual weight is 1(This can be varied based on words) then positive score is 5.

**Negative**: Check number of negative tokens. If there are 2 negative tokens and each token individual weight is 1(This can be varied based on words) then negative score is 2.

**Exclamatory**: Check number of exclamatory tokens. If there are 1 positive token and each token individual weight is 1(This can be varied based on words) and exclamatory token weight is 2 then total exclamatory count would be 1\*2 i.e. 2. For exclamatory before negative token we will negate by (2\*1).

**Inverters**: Check number of inverter tokens. If there are 1 inverter tokens before positive word then inverter value is -1 and if there are 2 inverter tokens before two negative words then total inverter count becomes 2. So the inverter count would be +1.

So that total count would be

Exclamatory + Positive - Negative + Inverters = sentimental\_factor.

In above case

sentimental\_factor. = 2+ 5 – 2 +1 = 8

**Abusive**: If there are any abusive words occur in token then we tag the review as abusive.

**Sample Data:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Entity Id** | **Review Text** | **Positive Count** | **Negative Count** | **Exclamatory** | **Inverter** | **Sentiment\_Factor** | **Abusive flag** |
| 123 | 234 | This is awesome. I really liked it. | Words:  Awesome, really  Count=2 | Count = 0 | Words: Really  Count=2 | Count=0 | 2+2-0=4 | No |
| 111 | 222 | This is fukcin waste of money. Utter Garbage.  I felt good after reviewing it. | Words:  Good  Count=1 | Words:  Waste, Garbage  Count=-2 | Words:  Utter  Count =  -2 | Count=0 | 1-2-2+0=-3 | Words:  Fukcin  Yes |
| 222 | 444 | This doesn’t work. Don’t buy. | Words:  Work, buy  Count=2 | Count=0 | Count =0 | Words:  Doenst,don’t both before positive words.  Count = -2 | 2-0+0-2=0 | No |

**Sentiment Analysis:**

Super Happy = sentimental\_factor > 5

Happy = sentimental\_factor > 2 and sentimental\_factor <=5

Neutral: sentimental\_factor =0 and -1

Sad: sentimental\_factor < -1 and sentimental\_factor > -5

Super Sad: sentimental\_factor < = -5