MINI PROJECT



**Name :- Ayush Raj**

**Section :- CE**

**Roll no. :- 20**

**University Roll no :- 2017317**

**Sem :- 03**

**Problem Statement:-**

Sentimental Analysis using Python to determine the polarity and nature of tweets data and classify them as Positive , Neutral or Negative.

**Introduction :-**

[Sentimental Analysis](https://www.paralleldots.com/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 (whether its positive ,neutral or negative). This is an effective tool that helps companies identify loopholes and rectify them as well as boom their product in their respective excellence fields.

With the recent advances in deep learning, the ability of algorithms to analyse text has improved considerably. Creative use of advanced artificial intelligence techniques can be an effective tool for doing in-depth research. We believe it is important to classify incoming customer conversation about a brand based on following lines:

1. Key aspects of a brand’s product and service that customers care about.
2. Users’ underlying intentions and reactions concerning those aspects.

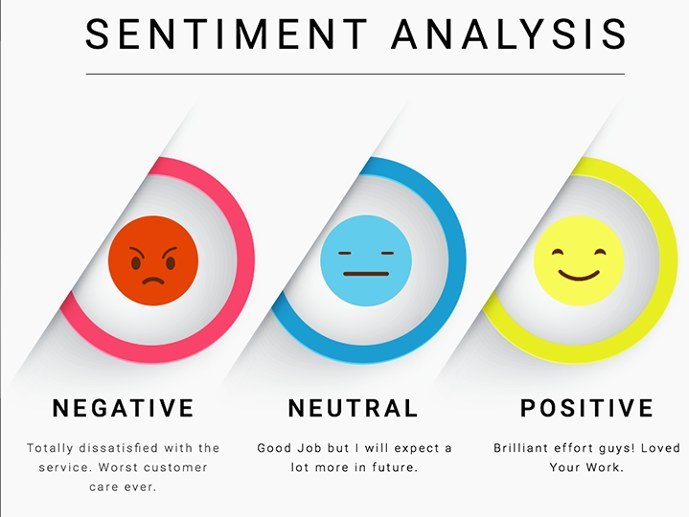


Fig. 1 :-Sentimental Analysis depiction

*https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.analyticsvidhya.com%2Fblog%2F2021%2F08%2Fsentiment-analysis-using-bidirectional-stacked-lstm%2F&psig=AOvVaw2B\_HjNBbVYJP2ayz5dWOvs&ust=1645689682991000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCJDKqvitlfYCFQAAAAAdAAAAABAD*

**Motivation behind doing sentiment analysis :-**

Out of different projects in Semester 3 , I found out this project quite appealing. The topic covers an important aspect – THE USER REVIEW AND OVERALL GROWTH OF A PRODUCT/SERVICE. Public Relations are quite important considering the success and failure of something . Moreover , on a technical basis I used it to learn some basic Python too to enhance my language skill in one more domain . I would also like to continue this further (if possible) so I can get deep into the concept of Machine Learning and Deep Learning and depict it on a larger basis. I am quite thankful to Mr. Vijay Singh, Assistant Professor , Dept. of CSE for his utmost support and guidance. I hope to continue upon this I near future so I can build several project under him and this language.

**Language/Tools used :-**

As it is not a too advanced or detailed project and basically a stepping stone into deep concepts , I used

* Python and related libraries.
* Jupyter Notebook to execute code.
* Data of around 31k tweets as to perform analysis on.

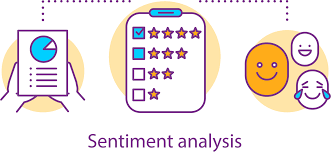


Fig 2:- Sentimental Analysis – Data analysis and information

https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.analyticsvidhya.com%2Fblog%2F2020%2F11%2Ffine-grained-sentiment-analysis-of-smartphone-review%2F&psig=AOvVaw15gZpXDUdfAKH2iYYCDjdW&ust=1645701615593000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCMijybPalfYCFQAAAAAdAAAAABAD

**HOW SENTIMENTAL ANALYSIS WORKS**:-

Lets take an example :-

*The Book was awesome !*

Step 1:- Tokenization.

Tokenization basically involves splitting up of the entire paragraph/review/rating into sub categories that are used later.

Eg :- The above sentence would be broken into fragments :

* *The*
* *Book*
* *Was*
* *Awesome*
* *!*

Step 2:- Cleaning up the Data.

Removing special characters so that only the core words are left to be examined on.

Eg :- ! would be removed and we will be left with *The , Book , was , Awesome .*

Step 3:- Removing the Stop Words.

Popular conjunctions /interjections are removed so that the words that add up sense are left with us.

Eg :- *The* and *was* are have no contribution and we will be left with *Book* and *Awesome*.

Step 4:- Classification and Scoring.

With the remaining words we classify them under respective categories. For positive we do + , for neutral we keep a 0 and for negative we have a -1.

Eg :- *Awesome* is a positive word (+1) and *Book* is a neutral word (0).

**THIS IS THE PLACE TO APPLY SUPERVISED ALGORITHMS**.

We can Train our model with some words or Lexicons, and test it on the analysing statements.

Step 5:- Calculation.

Now considering the statements , we calculate the overall polarity of the statement to see whether its positive , neutral or negative.

Eg :- Our overall score was +1(Awesome)+0(Book)=+1 …so it was a POSITIVE STATEMENT/Review.

**Applications of sentiment analysis :-**

Sentiment analysis tools can be used by organizations for a variety of applications, including:

* Identifying [brand](https://whatis.techtarget.com/definition/brand) awareness, reputation and popularity at a specific moment or over time.
* Tracking consumer reception of new products or features.
* Evaluating the success of a marketing campaign.
* Pinpointing the target audience or [demographics](https://whatis.techtarget.com/definition/demographic).
* Collecting customer feedback from social media, websites or online forms.
* Conducting market research.
* Categorizing customer service requests.
* Movie and book recommendation.

**Conclusion :-**

We have successfully analyzed sentiments of 31k tweets with an accuracy of 0.94.