

# Sentiment Analysis of Social Media Data for Efficient Opinion Mining

Srinidhi B S  
Research Scholar  
Jain University, Bangalore

Dr. Suchithra R  
Jain University  
Bangalore, India

## 1. INTRODUCTION :

Opinion mining is gaining success these days, because of its application in various fields. Mainly it gives us an estimate of what is happening regarding an entity may it be product or a movie or spanning across any social activities. Social media is a hub of data where it produces an excellent opportunity for researchers to analyze the data that are tweeted or posted across online.

These tweets are analyzed on various grounds, mainly **text analysis** is done on these to get the opinion on the trending topics across internet. Opinions are collected and tapped as big data, there are lot of software architectures to hold and analyze the big data with a minimum computational power too. This feature has made the problem interesting and reachable to the larger set of the researchers for various kind of analytics. These analytics help various companies to get the opinion at the real time for any entity of their choice. There are supervised and unsupervised learning algorithm implemented across, for the text analysis to detect the sentiment.

The research problem of interest here is more than the text analysis, we just speak of the text analysis and the efficiency at which we can analyze the opinion. While doing this we do forget the various other factors too. Importantly always while conversing before a person or a person expressing before the social media, it the emotional quotient which plays an important role for the wordings that he will be using in the comments or the tweets. And it is an important task for the researchers to know the emotional quotient to determine sentiment analysis effectively for further conclusions on the sentiment data.

### *Literature Survey*

Review of the literature for this work have spanned across the papers entitling from sentiment analysis procedures and big data and the Emotional quotient factor.

Opinion mining for reputation evaluation on unstructured Big Data has the comparison methods for the various machine learning algorithms on the sentiment analysis. It also reviews in detail about the sentiment granularity at which it can be performed, for a sentiment to be classified positive, negative and neutral what are the classifiers etc.. have been discussed

The Paper Emotional intelligence and its relation to everyday behavior discuss how it matters for the EI in everyday life and how the decision can depend and how a person ability revolves around it.

Paper Big Data Development, Challenges and opportunities," Lean Big Data, introduces us to the basics of the big data world, implementations and the challenges in the implementation of the real time huge data. There are whole lot of references made. I have given a short detail of what superficially was required, picking up one in each of the category.

## 2. OBJECTIVE AND SCOPE:

Objective for this research problem is to find an efficient way to implement the factors that can improve the sentiment analysis. Sentiment analysis is the one which is the trending these days because of its application in the academics and industry. People do the sentiment analysis to determine the actual value for an entity and to get a real-time feedback about the entity. This helps business to get hold on the market and make their future decisions intact. With this respect, the objective of this research is to find an efficient way to implement this analysis so that we make the decisions in a proper manner. This makes the research problem interesting for introducing new factors which makes the sentiment analysis efficient.

## 3. MOTIVATION :

When a thought about the sentiment analysis is given, we think only for the textual analysis. But in real time when we considering a review process, we just do not consider only one or two comments but we make a detailed study and then come to conclusions. The same can be considered for the sentiment analysis also, why to just make a textual analysis with what so ever accuracy but when we effectively prove that this is worth for further any data analysis. We often find heavier implementation methods to prove the accuracy of the text analysis. We use neural network and machine learning algorithms for an efficient understanding of the tweet or the comment, but the other factor more than the text analysis is important for an emotion detection is the motivation behind the thought of this work

## 4. DESCRIPTION OF RESEARCH WORK

Emotional Quotient refers to one's ability to perceive, control and evaluate emotions. This is an important factor which determines how a person will react to the situation and will he present himself for a situation. This is an significant factor since it is the attempt that an individual makes on the social media which he tells out in a louder

voice to the world about his or her feelings regarding an entity. It is always the Emotional intelligence that plays a vital role in the usage of the words as well as his/her behavior socially. This role of the text analyzers for the sentiment data is to pick up the tweets or comments made on the social network and analyze for the ratings of any product. Usage of language is so important for an analyzer so that, it could predict the right mood or rating for a particular entity. And the mood of the person depends on solely on the satisfaction/disgust/average feeling that he has experienced with the entity. Here is the actual need to look at the emotional quotient which is almost neglected for any sentiment analysis.

#### *Methodology of Implementation*

While we detail the study, we look at the pattern of the usage of the words among the various Emotional Quotient (EQ) people. Persons with different EQ are tested for their reactions on an entity, below would be the attached results of the usage of the word pattern for different EQ people.

Initially a study is being conducted on the person to judge their EQ factor. This is just a test given with a questionnaire with a set of questions having to judge their emotionally ability to react in various situations. This is a sample set to determine the EQ of a person. This answers would be a kind of a rating factor from 1 to 5. A mean average is calculated at the end to give an emotional score for a person. This score will determine the quotient of a person.

#### ***The Training set of questions are here as follows:***

1. In my group of friends, I am generally aware of how each person feels about the other people in our social circle.
2. When I am upset, I can usually pinpoint exactly why I am distressed.
3. While there are some things that I would like to change, I generally like who I am.
4. When I make mistakes, I often berate and criticize myself and my abilities.
5. I feel uncomfortable in emotionally charged situations.
6. I tend to avoid confrontations. When I am involved in a confrontation, I become extremely anxious.
7. I am generally aloof and detached until I really get to know a person.
8. I tend to overreact to minor problems.
9. I feel confident about my own skills, talents, and abilities.

10. I would describe myself as a good judge of character.
  11. During heated argument I tend to fight more.
  12. When taking important decision, I am follow my instinct
  13. I can easily make friends and get along with them
  14. When a project of my hard work is not recognized properly I get worried more
  15. One of your friend is in trouble, you would talk to him and convince about the situation?
- [Questions are taken from the <http://psychology.about.com/>]

At this level we can determine the behavior of the person socially and collect the tweets or the comments to determine the sentiment analysis. But when we consider the tweets from a user we cannot disclose the identity and it is unethical to analyze a social behavior of a user without his acknowledgement but for the analyzing the data in a huge set of tweets, it is also important to note a point that we will not reveal the identity at any cost.

For the training set purpose of the implementation, instead of real time tweets and comment collection we give two questions to the user and collect the way they feel about the entity and we analyze the sentiment along with the same replies

The Training questions are as follows:

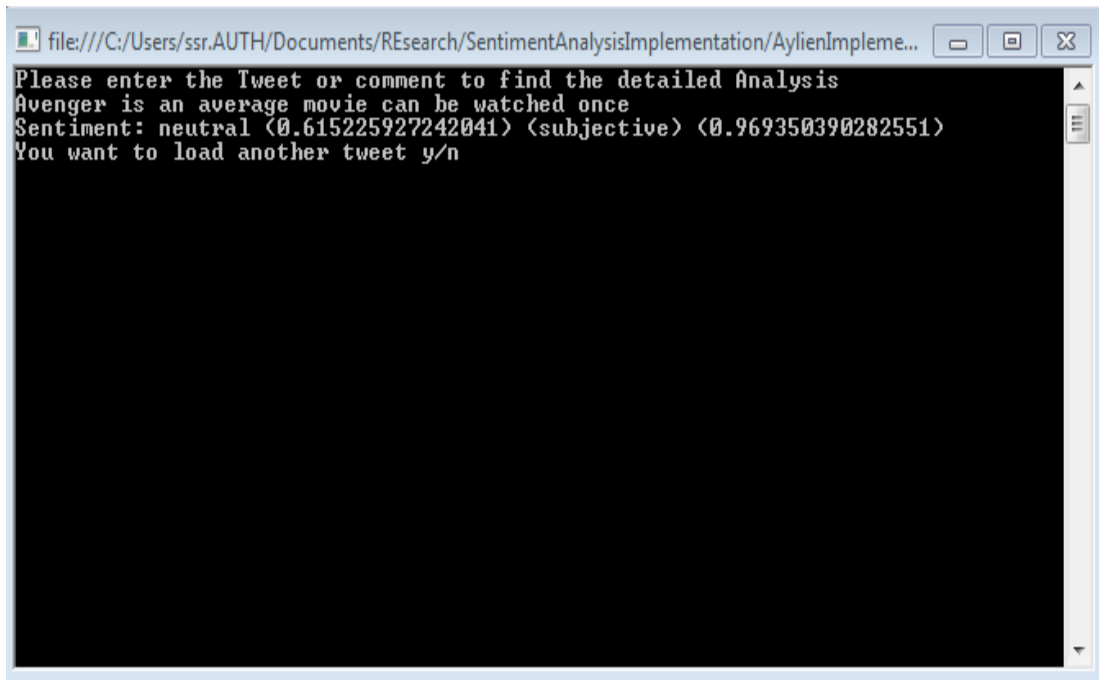
1. What is your expression about the recent movie that you had been to
2. What is your expression for the recent clothes you took

The answers are considered to be as the reaction socially and computed for the sentiment analysis on the same.

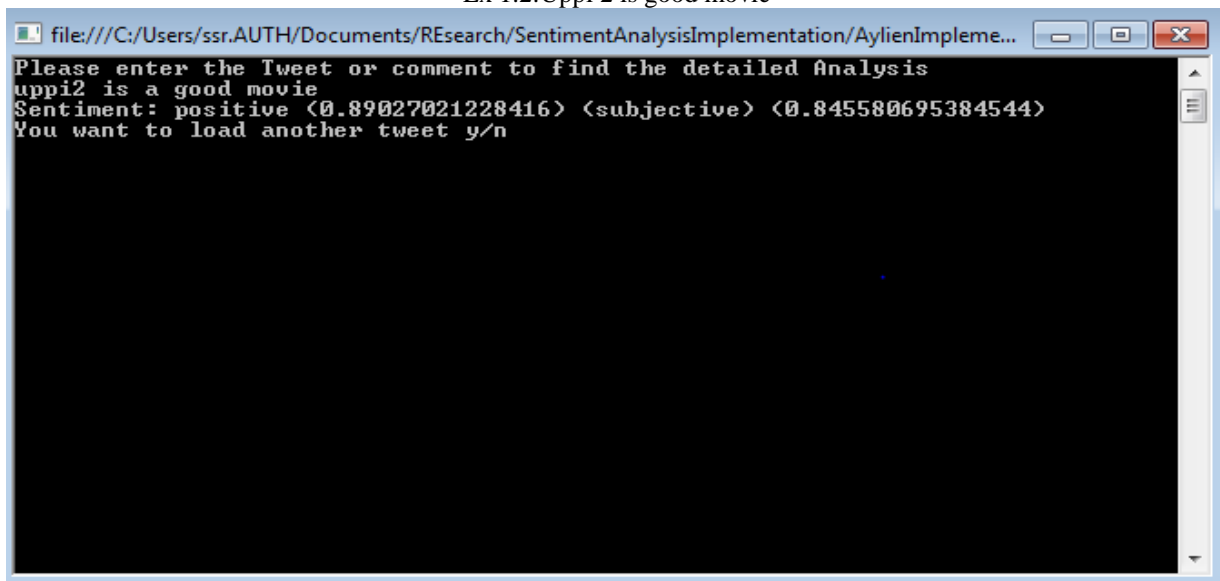
The above implementation uses the API provided but AYLIEN Software solutions to make the sentiment analysis, for the code is being developed in the dotnet framework to invoke the API. This API is implemented on a un-supervised learning pattern, and it has its own methodology of the implementations and learning pattern, We can just create a API key from the code side and invoke the rest service over HTTP protocol that it exposes.

The response from this is captured and it is de serialized to know each and every factor for the results of sentiment analysis. In our implementation here we calculate whether the sentiment is positive, negative or neutral we associate a score too telling how much positive a sentence is. A training set for the same discusses our problem of question in a sample manner

Ex 1.1: For an analysis on a tweet for the movie avengers: Avenger is an average movie can be watched once  
Overall Sentiment: Average. With a score of 0.6 and the sentence is subjective

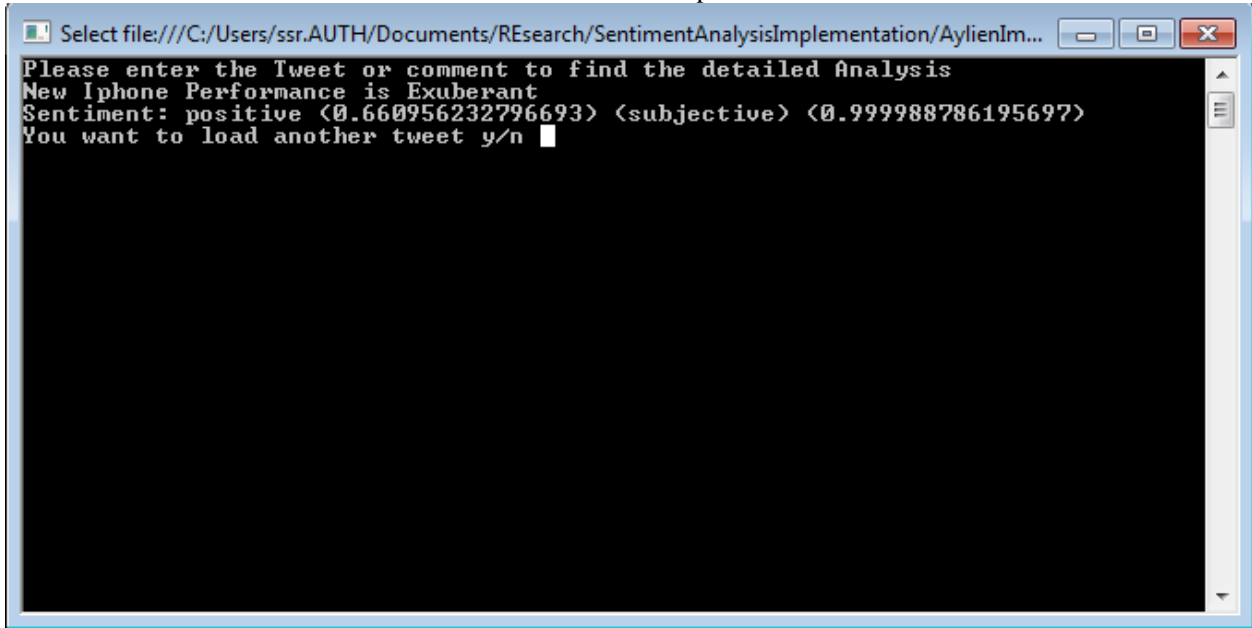


Ex 1.2:Uppi 2 is good movie

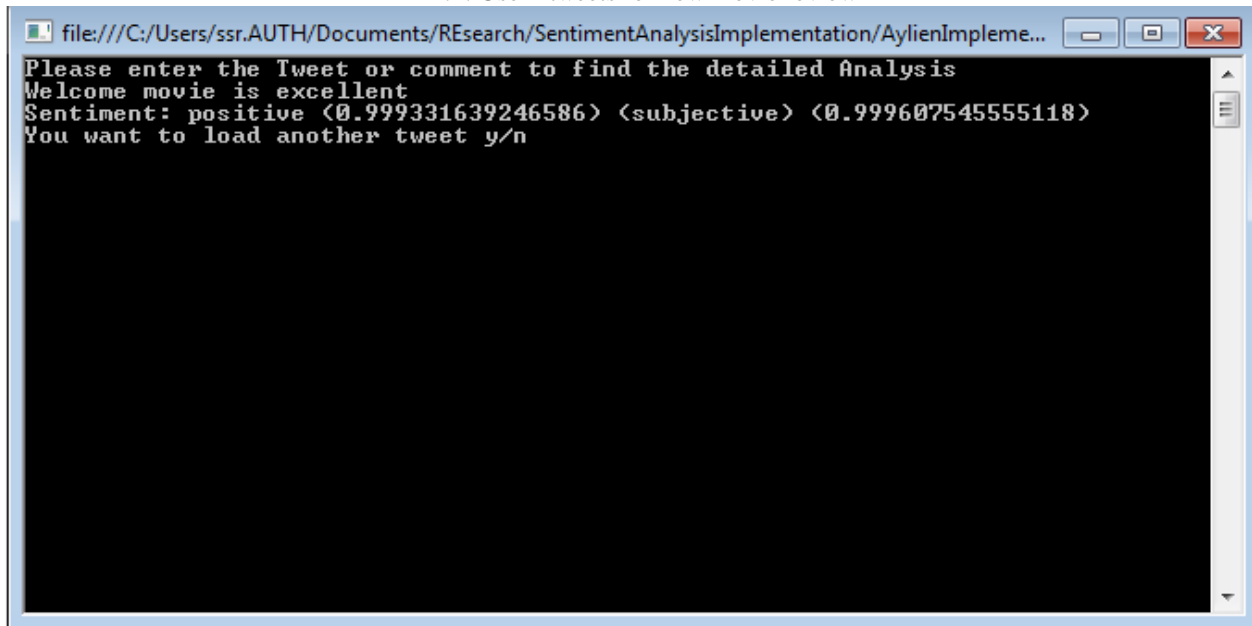


For this statement Sentiment is positive with 0.8 polarity  
Like the same we can categorize the tweets and comments on the Facebook and come out with a detailed analysis regarding the opinion of an entity.

Ex2.1:User1 tweets with the new iPhone performance to be exuberant

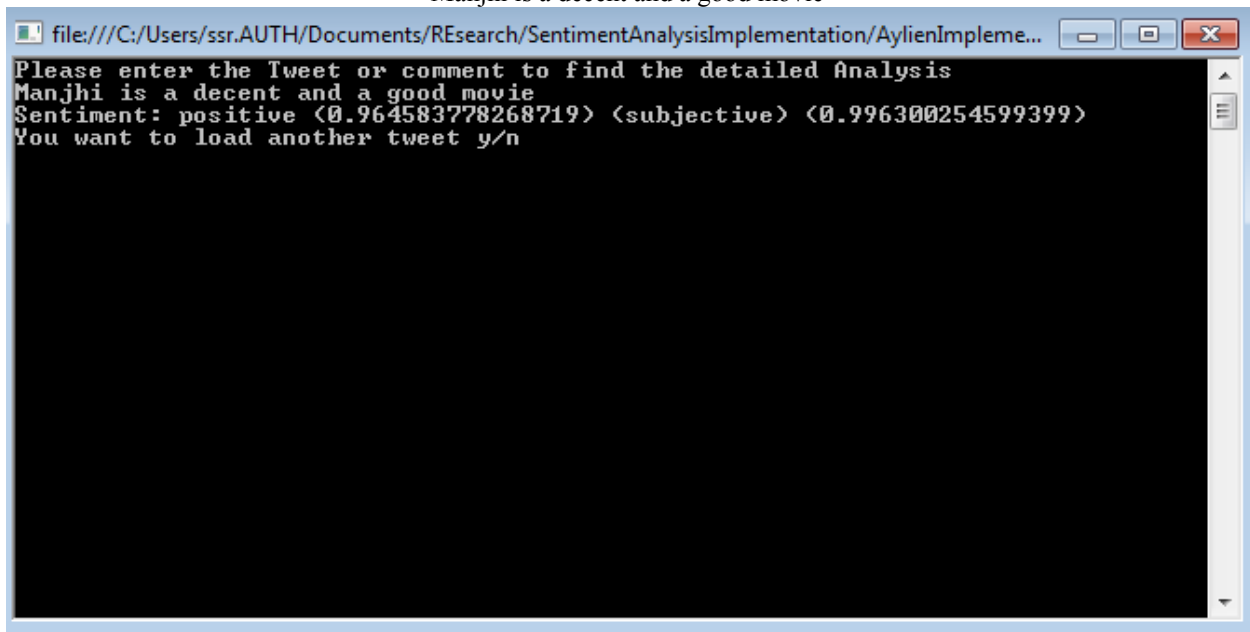


Ex2.2: User1 tweets for new movie review



When the phrases are used or either re phrased, the score of the sentiment either decreases or increases. As an example when we consider the Tweets of the User2 at almost same scenario for different movie review given by him, this is how he express the same

Ex 2.3: When user 2 express the review on movie Manjhi he mentions that as:  
 “Manjhi is a decent and a good movie “



## 5. CONCLUSIONS :

1. From the above results it is clear that the results can depend on the factor of the emotional quotient of person in predicting the mood
2. EQ of a person will determine the pattern of the usage of the language
3. Sentiment analysis with considering these factors will add the accuracy of the ratings that we give for an entity after analysis
4. From the attached excel sheet we can identify that a person with the average EQ would use a pattern of words which can generate a average sentiment score

## 6. REFERENCES :

1. “Big Data Development, Challenges and oppurtunities,” *Lean Big Data*, 24 February 2014.
2. “Social Network Analytics,” *LeanBigData*, 5 April 2014.
3. “Data Modelling For Big Data,” *CA TechnologieData*.
4. “Big Data Model or Not Model?,” *SQL Server*.
5. D. H. M. Gali Halevi, “Evolution of Big Data as a Research and Scientific Topic,” *Research Trends*, no. 30, September 2012.
6. C. K. Boyd D, “Critical Questions for Big data,” *Information, Communication & Society*, no. 15 (5), p. 662, 10 May 2012.
7. “Big data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey Global Institute.
8. B. D. & C. K, “Six Provocations for Big Data,” in *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, 2012.
9. G. M. & G. M. De Mauro A, “What is Bi Data? A Consensual Definition and a Review of Key Research: Topics,” in *4th International Conference on Integrated Information*, Madrid, 2014.
10. “Emerging Technologies: Big Data - A Horizon Scanning Research Paper,” The emerging technologies big data community of interest.
11. A. Rossano, “The Pragmatic Side of Using Big Data,” 2013.
12. <http://psychology.about.com/>
13. <http://docs.aylien.com/>
14. Twitter as a Corpus for Sentiment Analysis and Opinion Mining Alexander Pak, Patrick Paroubek
15. Emotional intelligence and its relation to everyday behaviour ,Marc A Brackett,John D Mayer, Rebecca M Warner
16. Opinion mining for reputation evaluation on unstructured Big Data by Uma Gaurav and Nandini