



International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 10 • Number 19 • 2017

An aggregated approach to Sentiment Analysis

Sanjib Kumar Sahu¹, Pankaj Kumar², D.P. Mohapatra³
and Rakesh Chandra Balabantaray⁴

¹ Dept. of Computer Science and Application, Utkal University, Bhubaneswar, Odisha, India,
Email: sahu_sanjib@rediffmail.com

² University School of Information and Communication Technology, Guru Gobind Singh Indraprastha University,
Delhi, India, Email: pankaj786067@gmail.com

³ Dept. of Computer Science and Engineering, NIT, Rourkela, Odisha, India, Email: durga@nitrkl.ac.in

⁴ Dept. of Computer Science and Engineering, IIIT, Bhubaneswar, Odisha, India, Email: rakesh@iiit-bh.ac.in

Abstract: Sentiment Analysis also called as opinion mining is one of the major task in NLP (Natural language Processing) and has gain much attention in recent years especially in social media platforms. In this paper, we aim to tackle the problem to analyse the sentiments on multiple social media platforms which is one of the fundamental problems in social media analytics. We have used data from Twitter, Facebook and YouTube for iPhone 7 mobile device to analyse the audience sentiment for this product, and further analyse the result by comparing the sentiment between these platforms and presenting a cumulative sentiment study for respective product. At last, we also give insight into our future work on sentiment analysis.

Keywords: Sentiment Analysis, Supervised Learning, AFFIN's List, R, S QL, Information Retrieval

1. INTRODUCTION

Sentiment Analysis plays an important role for analysing the sentiments of user over different social media platforms and producing a meaningful result which will help people or organizations in making important decision for their respective product and services in future. In recent years, digital data has become very easy because of fast development of the Internet and social media platforms, where people share their views and thoughts about selective issue, topic, product or anything else in general. Now -a-day's people are sharing their views not only over one social media platform but over multiple plat forms like Facebook, Twitter, YouTube and many more. So there is a strong possibility that a user which is present on one social platform may not be present on other platform, therefore for analysing the sentiment for a particular topic, we required to take in account multiple platforms and present a cumulative study which would be more accurate rather than analysing a single social media platform. In this study we are taking in account three of the major social media platforms i.e. Facebook, Twitter and YouTube for analysing sentiments of user review for iPhone 7 mobile device.

Social Media: A common online platform where people share, comment, like and show interest over a specific or generic topic.

Sentiments: A sentiment is a view or opinion about something. A way of expressing you thoughts in response to an event in past, present or in future.

Sentiment Analysis: Study of analysing sentiment using modern natural language processing techniques and concluding the result in form of positive, negative and neutral sentiments. In some cases it is hard to determine, that a sentiment belongs to which of the three mentioned categories, for these case we used different machine learning techniques.

The rest of the paper is organized as follows. Related Work are presented in section II. Proposed algorithm are explained in section III. Experimental results are presented in section IV. Concluding remarks are given in section V.

2. RELATED WORK

Sentiment Analysis is an old concept in area of natural language processing and social interaction, which has been studied by many researcher in past to analyse the sentiments of user. Social media is growing at very fast pace because now-a-days social interaction is recorded over social media platforms like Facebook, Twitter, YouTube and others in real time. Sentiments like happiness, anxiety, excited, opinions, beliefs, reviews, comments, explanations, sadness, courtesy and many more are shared among various post and tweets over a selective topic of interest. As a part of my study some of the related works are presented below:

As in Paper [1], author has conducted a study of sentiment analysis over social media texts for Twitter tweets by analysing the sentiments as positive, negative and neutral using unigram and bigram approach. As in Paper [2], author has described how sentiment analysis can be used to analyse document, entity and sentence level which is helpful for disaster relief during emergency situations. In paper [3], author has compared the sentiment analysis for online movie reviews from IMDB and Twitter tweets, based on Naïve Bayes approach author determines the accuracy of sentiment among these datasets. In paper [4], author has presented an overview of most recent capabilities in social media analysis, which is further categories in three major do mains: content analysis, group/network analysis and prediction of real world-events. In paper [5], they explained the importance of sentiment analysis and explain various methods of extracting relevant information in different data format like HTML, JSON and many more for analysing sentiments and opinions.

In all the paper referenced above, we found that the nature of analysis conducted so far is related to single social media platform, and in some cases more than one. Now-a-days there are multiple social media platforms and people are more active online rather than offline approach for sharing information's and opinions. In this paper we are presenting how we can analyse the sentiments over multiple social media platforms and analyse the sentiments by comparing the result among multiple platform for a product, services, topics etc. based on the proposed algorithm explained in section III, we can expand this study to n number of social media platforms for more accurate and better results.

3. PROPOSED ALGORITHM

3.1. Sentiment Analysis Algorithm

The algorithm is designed to analyse the data collected from different social media platforms and perform sentiment analysis on each statement in the data collections, based on which it will provide the sentiments score with respect to each statement in the collection. A simple flow chart is shown below:

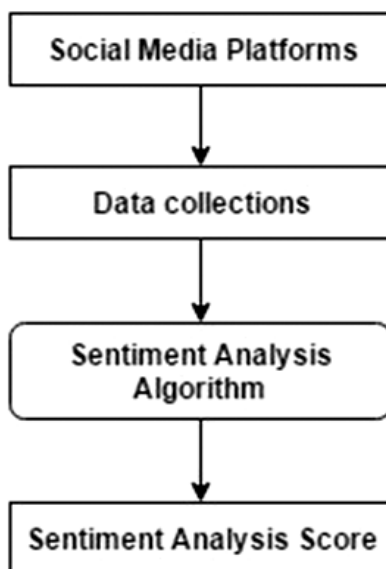


Figure 1: Sentiment Analysis Algorithm Flow Chart

On the basis of such considerations, the algorithm reads data from multiple social media platforms, execute the sentiment analysis algorithm to generate the sentiment scores, compares the scores and finally present the cumulative analysis result.

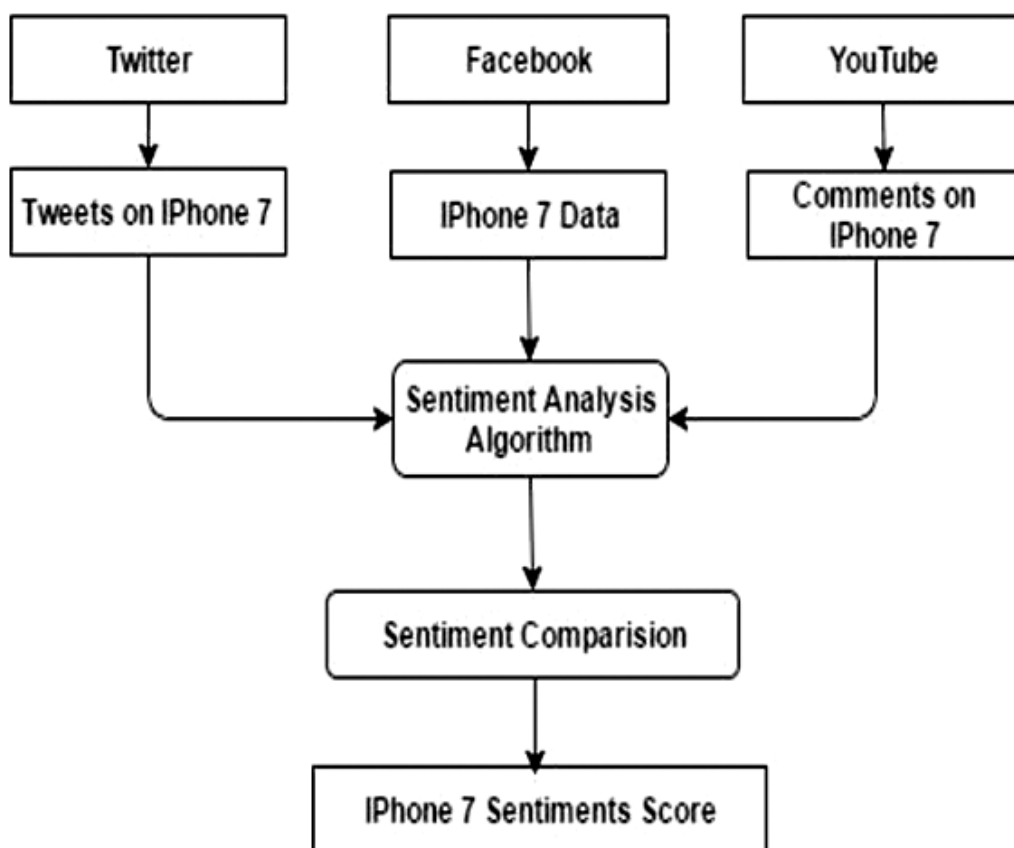


Figure 2: Sentiment Analysis on T witter, Facebook and YouTube

Once we got the sentiment scores from Sentiment Analysis Algorithm from different social media platforms we take the average of positive, negative and neutral sentiments for respective cumulative sentiment result. So, let assume PS (t), PS (f), and PS (y) stands as follows:

PS (t): Positive Sentiment score for Twitter

PS (f): Positive Sentiment score for Facebook

PS (y): Positive Sentiment score for YouTube

PS (x): Final Sentiment Score

$$PS(x): (PS(t) + PS(f) + PS(y)) / 3 \tag{1}$$

Similarly for Negative (N (x)) and Neutral (Nu (x)) sentiments we can calculate the cumulative sentiment score. Here 3 stands for total number of social media platform we are taking in account. In compare to positive and negative sentiments neutral sentiments doesn't strongly participate in decision making process so we often try to avoid the neutral scores in sentiment analysis but in this paper we have consider all of them.

4. EXPERIMENT AND RESULT

Using supervised learning techniques for machine learning we are using, dataset which contains 100 records on iPhone 7 from each social media platforms, so we have 100 tweets from Twitter, 100 comments/post from Facebook and 100 comments on iPhone 7 from YouTube video, are used to perform this experiment. For data collection with these social media sites we have follow the following process:

1. Twitter API to get the tweets from twitter.
2. Facebook Graph API and Netvizz v1.4 app to stream the live comment and posts
3. YouTube API to get the comments for a particular video in JSON format.

For analysis purpose we have used R programming language and related packages, SQL for aggregate score analysis, AFFIN's dictionary to assign sentiment scores for each word in the statement. Once the data collection is ready we executes the algorithm which will reads records from data collection file and stores into the defined local variable. After algorithm breaks each statement into multiple words, based on AFFIN's dictionary we assign scores to each word in the statement ranging from -5 to +5 as follows:

Table 1
Sentiment Score based on AFFIN's Dictionary

<i>Sentiments</i>	<i>Scores</i>
Very Negative	-5 to -4
Negative	-3 to -1
Positive	+1 to +3
Very Positive	+4 to +5

Once the scores has been assigned to each word of the statement of respective data collections , we use SQL for combining the intermediate results and then apply aggregate function to join the scores together for respective data collections. The output for all the data collection has been shown below

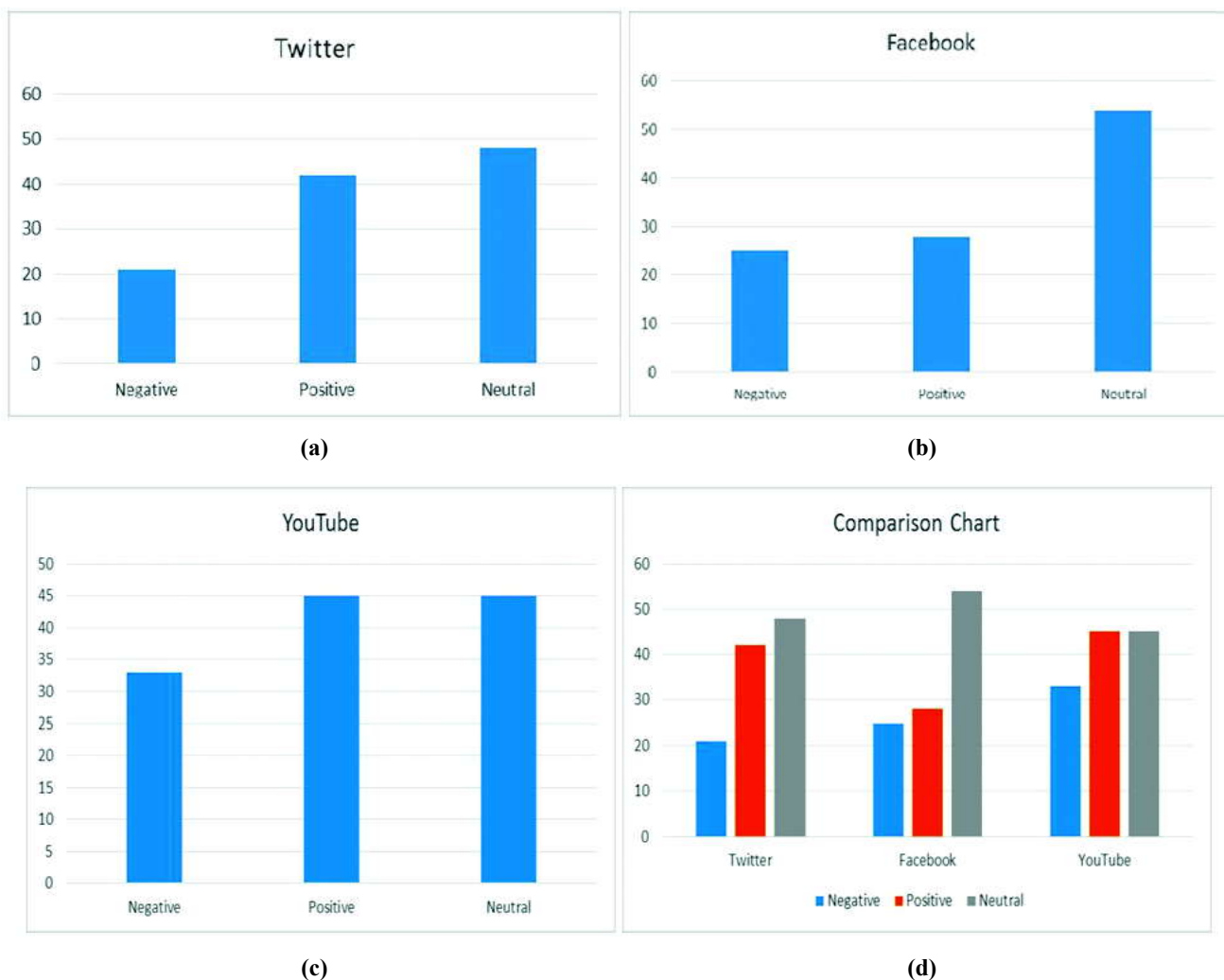


Figure 3: (a) T witter Sentiments (b) Facebook Sentiments (c) YouTube Sentiments (d) Cumulative Sentiments of all 3 social media

Table 2
 Experiment Result
 for iPhone 7

	<i>Negative Sentiments</i>	<i>Positive Sentiments</i>	<i>Neutral Sentiments</i>
Twitter	21	42	48
Facebook	25	28	54
YouTube	33	45	45

Table 2 show the positive, negative and neutral sentiments for each of the social media platform. As here we can see that response is different based on different social media which give us broader perspective of how people feel about the product.

Also, we can analyse the sentiments in a cumulative form as shown below, the score inside the table shown below is generated by formula mentioned in (1) for all 3 sentiments (positive, negative and neutral)

Table 3
Experiment Result for iPhone 7

	Negative	Positive	Neutral
Cumulative Sentiments	26	38	49

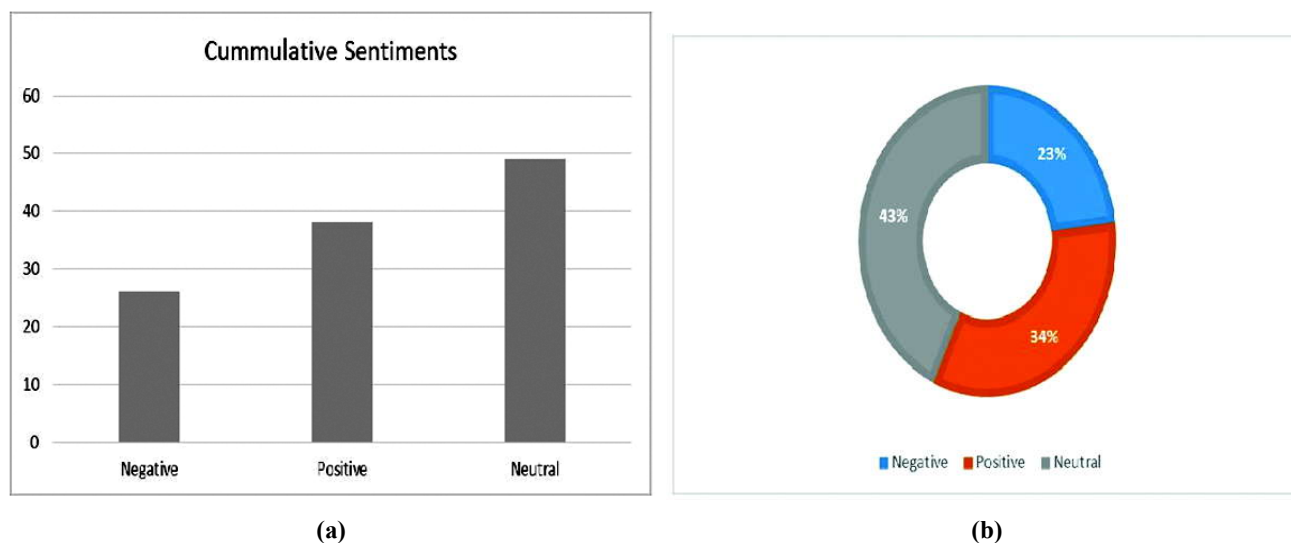


Figure 4: (a) Cumulative Sentiments (b) Percentage (%) wise distribution

5. CONCLUSION

In this work we have presented a new approach to perform sentiment analysis on multiple social media platforms. Unlike standard method of analysing the user sentiments over a single social media platform this method will provide more accurate sentiment score with respect to the product or services we are trying to analyse. Not only this, as we increased the number of social media platform for a specific topic we will be getting better results. In our future study we will be focusing more on analysing the statements having multiple meanings which is hard to analyse by normal sentiment algorithms, also our focus would be analysing the sentiments which are more intensity based rather than only focusing on negative, positive and neutral sentiments.

REFERENCES

- [1] Alexandra Balahur, *Sentiment Analysis in Social Media Texts*, European Commission Joint Research Centre Vie E. Fermi 2749 -21027 Ispra (VA)
- [2] Ghazaleh Beigi, Xia Hu, Ross Maciejewski and Huan Liu, *An Overview of Sentiment Analysis in Social Media and its Applications in Disaster Relief*, 2014
- [3] Rajni Singh, Rajdeep Kaur, *Sentiment Analysis on Social Media and Online Review*, *International Journal of Computer Applications* (0975–8887) Volume 121 – No.20, July 2015
- [4] ZN Gastelum, KM Whattam, *State-of-the-Art of Social Media Analytics Research*, Pacific Northwest National Laboratory operated by Battelle for the U.S Department Of Engergy under Contract DE -AC05-76RL01830, Jan 2013
- [5] Bogdan Batrinca, Philip C. Treleaven, *Social media analytics: a survey of techniques, tools and platforms*, *AI & Soc* (2015) 30:89–116 DOI 10.1007/s00146-014-0549-4, 26 July 2014
- [6] ZN Gastelum, KM Whattam, *State-of-the-Art of Social Media Analytics Research*, Research Paper, Pacific NortWest Laboratory, 2004.

- [7] Andy Bromberg, *Sentiment Analysis in R*, <http://andybromberg.com/sentiment-analysis>.
- [8] Xing Fang, Justin Zhan, *Sentiment analysis using product review data*, *J. Journal of Big Data*, Dec 2015
- [9] T witter API, <https://dev.twitter.com/overview/api>
- [10] Facebook Netvizz v1.4, <https://apps.facebook.com/netvizz/>
- [11] YouTube API, <https://console.developers.google.com/>
- [12] G Vinodhini, RM Chandrasekaran, *Sentiment Analysis and Opinion Mining: A Survey*, Volume 2, Issue 6 ISSN: 2277 128X, June 2012