Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 6, June -2017

Comparative Study of Opinion Mining Tools

Kanika Garg[1*], Goonjan Jain[2]

[1]School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India;
[2]Department of Applied Mathematics, Delhi Technological University, New Delhi, India

Abstract:- *Now-a-days, Opinion Mining is considered as one of the most demanding field of natural language research. Also called as Sentiment analysis aims to capture sentiments, emotions, opinions and attitude of speaker or writer. These emotions, sentiments etc. can be made towards other person, any product or service, any topic of interest such as politics, movies and games etc. Online blogs, discussion forums has provided people with a platform where they can share their views very easily. To mine these views various opinion mining tools are available online which can be used to find out the polarity of the text. So, this paper focusses on these tools, their capabilities leading to conclusions about them.*

Keywords: *Opinion Mining, Sentiment Analysis, Opendover, SentiStrength, Lexalytics, Sentimetrix, Text-Processing*

1. Introduction

Opinion Mining is a field of Natural Language Processing and Information Retrieval which mainly focusses on subjective information. Information generally contains either factual information called as objectivity or different expressions called as subjectivity. Subjective information is an important part for the research community as it contains views, sentiments and opinions of different individuals.

Over the last decade, there is an extensive use of Web 2.0. This has provided users with a platform to share their ideas globally. Various online sites such as web blogs, discussion forums and various social mediasites such as Facebook, twitter etc. and online shopping sites with customer reviews such as amazon and flipkart etc. began to flourish. People started using these platforms to get opinion of various products and services before actually using it and also started sharing their views, experiences of the services and products they have already used on this global platform. Consumers are using this platform to share their grievances, people are using this to appreciate or criticize politicians, actors or anything they care about. This has led to increased volume of opinionated data available online.

Opinion mining offers many tools and methods to develop new applications that are able to deal with huge amount of information present on web. E.g. Recommender Systems[1] produce recommendations regarding services or products based on the reviews and different aspects that may provide sentimental information regarding them. Subjectivity detection[2] helps in identifying the subjective sentences that are useful for any organisation and are more valuable. Summarization[3] leads to summarize the opinion regarding certain product or person from different sources. Every organisation small or big, country's government wants to know public opinion regarding their candidates, products and services. So, there are plenty of applications that opinion mining offers. Some of the application areas are:

- Decision making
- Recommendation systems
- Business strategies
- Product analysis
- User Modelling

Understanding human feelings, emotions and opinion is not a simple task. It needs a lot of effort and mathematical modelling to achieve good results. Many companies and research groups are creating NLP tools for carrying out sentiment analysis on a global scale. These tools can be imparted into other systems or research work without being an expert of that field. Thus, this work is focussed on free sentiment analysis tools available online. In this work, detailed description of each tool is provided along with the main functionalities of these tools.

This work presents in-depth and complete information regarding many sentiment analysis tools. The remainder of the work is organised as follows: Section 2 presents the related work in the field of opinion mining. Section 3 presents the main characteristics of many freely available tools. Section 4 presents the comparative analysis of various discussed tools. Finally, Section 5 draws important conclusions.

2. Background

To understand sentient analysis, it is important for us to understand some basic terminologies. There is a dire need to understand the difference between various terms that has been used as synonyms of each other, e.g. Opinion Mining, Subjectivity detection and sentiment analysis [4].

2.1 Basic Terms

Any sentence we write or speak can be either subjective or objective. Subjective sentence is the one that has some polarity either positive or negative or contains some emotion[5]. Objective sentences have facts and which have genuinely happened.So, subjectivity detection is basically detecting subjective sentences from a pool of mix sentences. Only subjective sentences are of importance in respect to sentiment analysis.

Opinions are a viewpoint of a person or a community which may deals with subjective matters with no concrete conclusions[6]. It expresses views on an entity or on some features of that entity. Further, opinions can be comparative sentences that express the resemblance between entities having common features and aspect[7].

Emotions on the other hand are different from opinion and sentiments. Opinions are the belief that a person forms about some object but emotions are physiological process of internal feelings of a person. Emotions express state of a mind of person regarding any subject. Emotions have more impact on human mind than opinions. So, to analyse emotions is much more difficult task than opinion mining and appraisal theory. Lot of work is going on appraisal theory[8]as it is useful in sarcasm detection[9], detecting stressful events, detecting abuse[10] and other various unlawful activities that can happen can be detected using social web.

2.2. Levels of Sentiment Analysis

1. *Document Level:* An assumption is made here that a document contains an opinion about single entity and by single opinion holder. Therefore, it is most suitable for regular sentiment. But in the case of blogs and discussion forums, multiple entities are compared with similar properties and characteristics; hence document level analysis is not desirable in comparative opinions. Subjectivity/objectivity classification is very important in this type of classification. Some of the tools that we will discuss work at this level. E.g. Lexalytics[11], Opendover[12]

2. *Sentence Level:* In the sentence level sentiment analysis, the polarity of each sentence is calculated. Objective and subjective sentences must be found out. The subjective sentences contain opinion words which help in determining the sentiment about the entity. After which the polarity classification is done into positive and negative classes. SentiStrength[13] works at sentence level.

3. *Word/Phrase Level:* The phrases that contain sentiment words are found out and a phrase level classification is done. The contextual polarity of the phrase in which a word appears may be different from the word's prior polarity. Almost every tool works at word level.

3. Tools

This section summarizes the characteristics and features of various freely available online tools. Also, they offer easy access to perform experiments on our own data. The various tools are as follows:

3.1 Opendover

It is Java based service provided freely on web. The main functionality of this web service is to extract semantic information from content management systems, texts and blogs etc. It provides full semantic tagging, automatic detection of context dependent, automatic disambiguation of words related to a particular domain and context independent sentimental words. It is based on ontology. It has a vast knowledge base for sentiment bearing words and their relationships.

It classifies sentiments into three emotional groups: judgement, appreciation and emotion. Each senti-word is labelled with a polarity and its score. The polarity score is on the scale of [-9,9] and the polarities are positive, negative and neutral. Sample data and its output in shown in Fig.1.

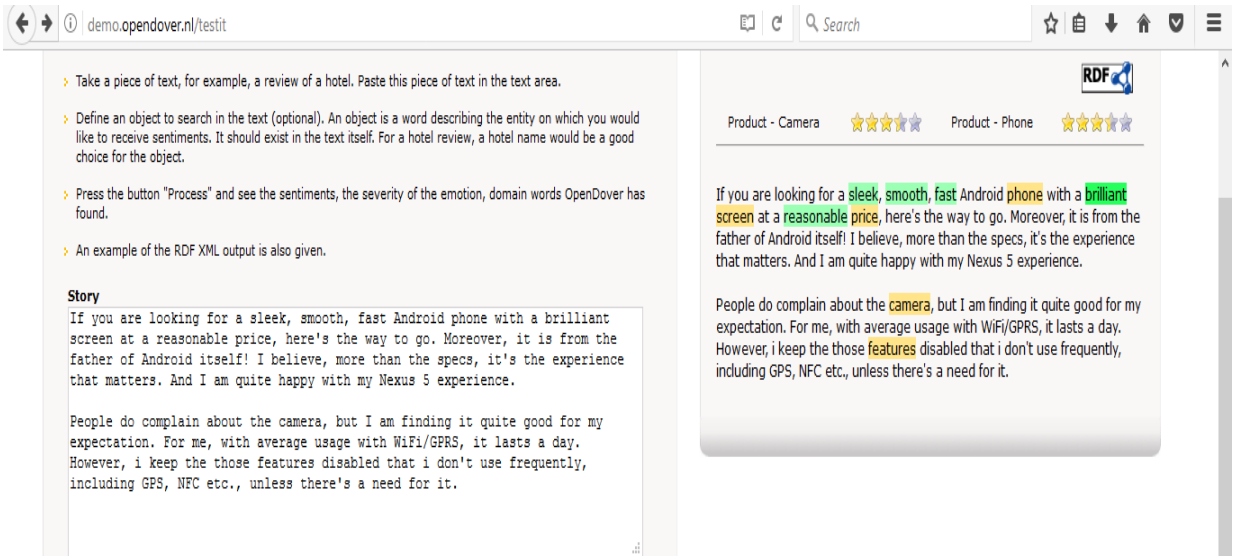

Fig.1: Sample amazon review and its output in Opendover

3.2 SentiStrength[14]

It is a sentiment analysis web service and tool created for short text. It classify text basically into positive or negative especially in English. Although it has been introduced for other languages as well. The various other languages it supports are Finnish, German, Dutch, Spanish, Italian and Russian.

This is also used in various social media analysis software suits, e.g. COSMOS which is a multiplatform open data analytics software, Mozdeh based on windows for Big data text analysis and Chorus data harvesting and visual analytics suite for Twitter.It takes small sentences to predict the sentiment. Here's a glimpse (Fig.2).

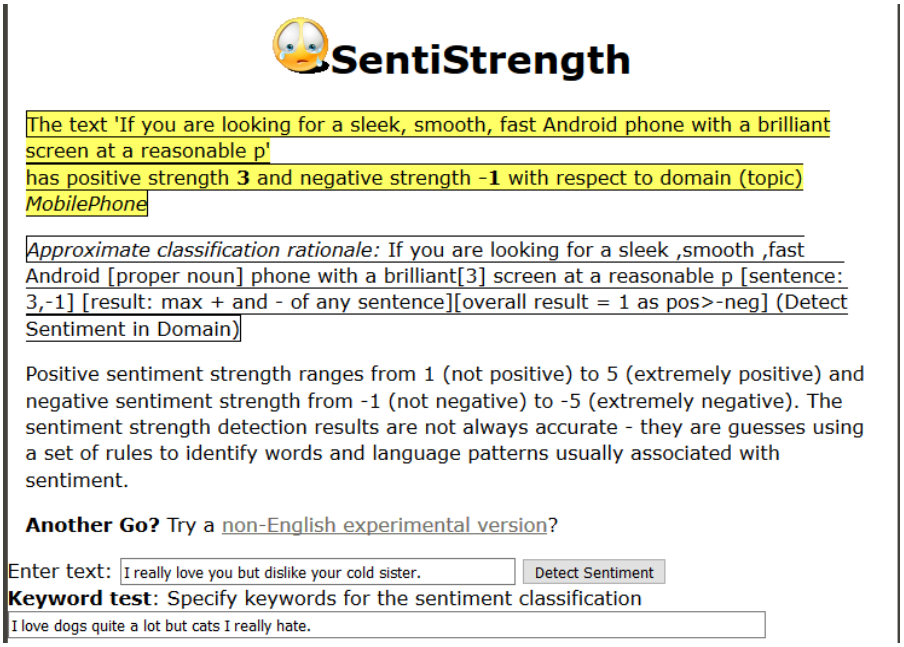

Fig.2: Sample amazon review and its output in SentiStrength

3.3 Lexalytics

This is document level software for sentiment analysis in the range [-1 1]. This software gives importance to the middle terms like 'but', 'if' etc. so as to calculate sentiments more accurately. This software enable to find sentiment about specific entities, people, products or services. It provides direct business solutions as well as API's to integrate in other's software's. It supports many languages like English, French, Arabic, German, Hebrew, Japanese, Korean, Italian etc. The figure below (Fig.3-5) shows the working of this software on a document containing text to be analysed.

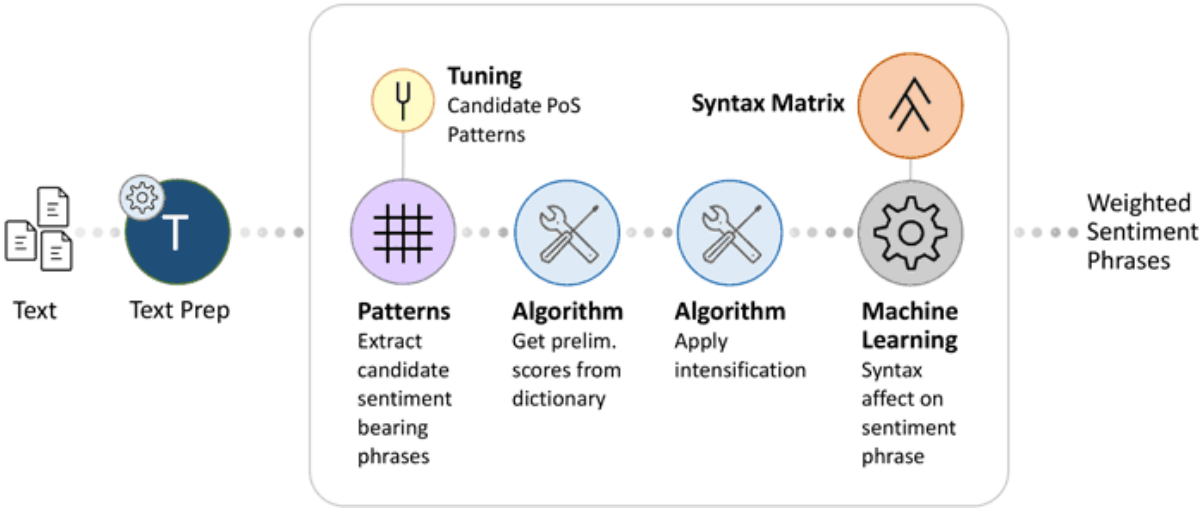

Fig.3: Work Flow of Lexalytics

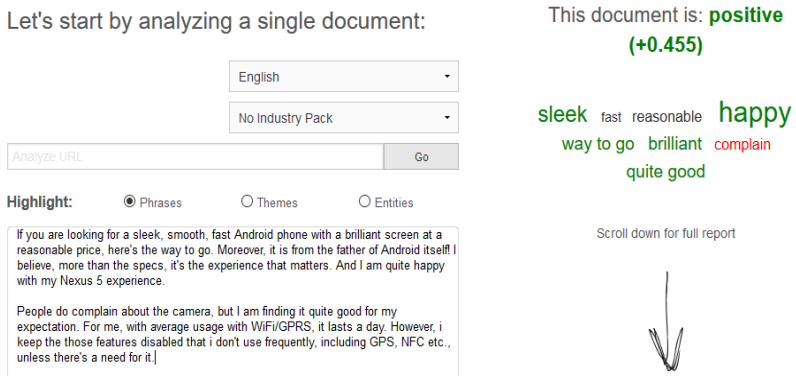

Fig.4: Overall output sentiment of sample amazon review using Lexalytics

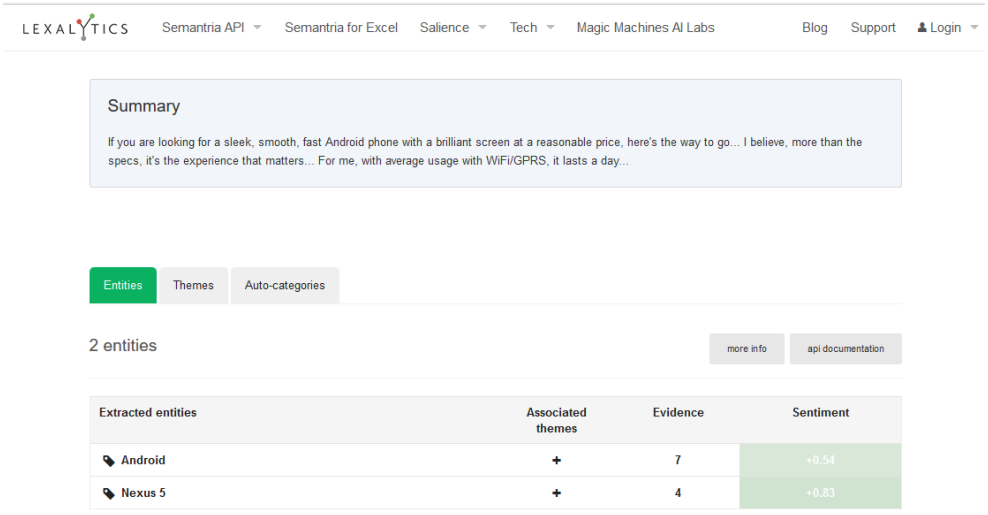

Fig.5: Summary and sentiments regarding associated entities of sample amazon review

3.4 Sentimetrix[15]

It is a commercial tool that provide social analytic solutions to various companies. Its main objective is to analyse consumer's opinion and expressions over products and services provided by such companies. Its expertise is in data analysis, language analytics, social network and social media analytics, computational social science, artificial intelligence, statistics and mathematics.

It provides various services related to sentiment analysis that are:

- SentiGrade: It's an API and a dashboard Web-based interface that allow customers to track sentiments from a given piece of text. It allows to compare changes in the sentiments from time to time.
- SenteMotion: It's an API and a dashboard Web-based interface that tracks intensity of different human emotions in any given document.
- SentiBility: It performs social media analytics to predict the stability of any government in a country based on information available on online sources.
- SentiSocial: It's a browser based dashboard that identify and display trend setter and amplifiers of opinion in real time.

Many major companies are using services from Senti-metrix like Forbes, TCS, The World Bank etc. This platform can be used via Gate[16] a Java based tool to perform NLP operations (Fig.6).

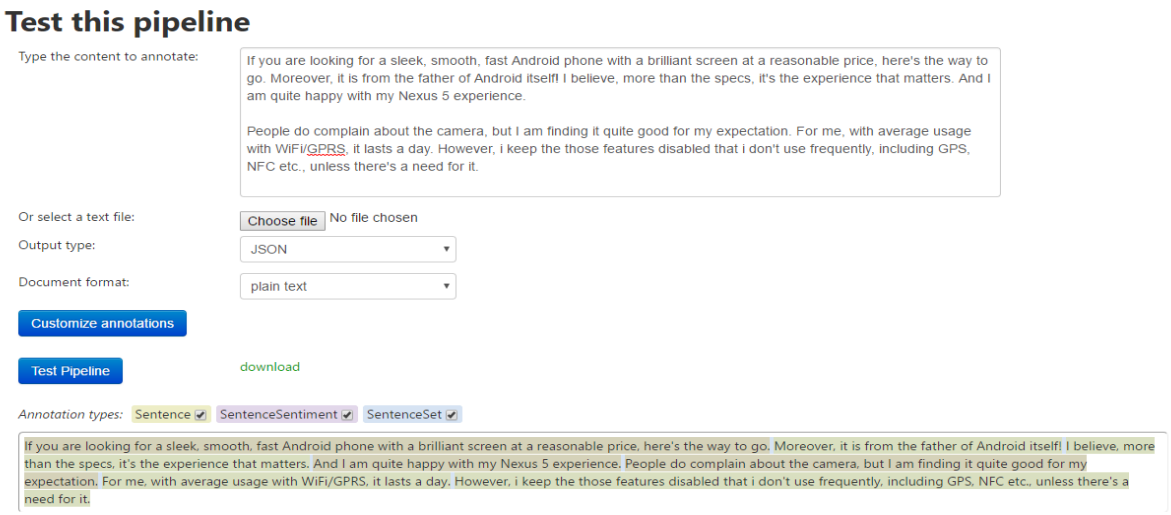

Fig.6: Sample amazon review in Sentimetrix

Output file as JSON: {"text":"If you are looking for a sleek, smooth, fast Android phone with a brilliant screen at a reasonable price, here's the way to go.Moreover, it is from the father of Android itself! I believe, more than the specs, it's the experience that matters. And I am quite happy with my Nexus 5 experience.\n\nPeople do complain about the camera, but I am finding it quite good for my expectation. For me, with average usage with WiFi/GPRS, it lasts a day. However, i keep the those features disabled that i don't use frequently, including GPS, NFC etc., unless there's a need for it.","entities":

{"SentenceSentiment":[{"indices":[0,127],"sentiment_string":"smooth","polarity":"positive","score":0.5,"sarcasm":"no","emotion":"good","rule":"SentenceSentiment","rule2":"SentiNoChange","holder":null},{"indices":[245,293],"polarity":"positive","score":1,"sarcasm":"no","emotion":"happy","rule":"XVerbSentiment","sentiment_string":"quite happy","entity_target":"I","rule2":"SentenceSentimentTarget","holder":null},{"indices":[295,382],"sentiment_string":"quitegood","polarity":"positive","score":1,"sarcasm":"no","emotion":"happy","rule":"SentenceSentiment","rule2":"SentiNoChange","holder":null}],"Sentence":[{"indices":[0,127],"sentiment":"positive","emotion":"good","secondPerson":"yes"},{"indices":[128,178]},{"indices":[179,244],"firstPerson":"yes"},{"indices":[245,293],"sentiment":"positive","emotion":"happy","firstPerson":"yes"},{"indices":[295,382],"sentiment":"positive","emotion":"happy","firstPerson":"yes"},{"indices":[383,441]},{"indices":[442,569],"firstPerson":"yes","directive":"imperative"}],"SentenceSet":[{"indices":[0,569],"score":0.8333333,"score_std_dev":0.28867513,"polarity":"positive"}]}}

3.5 Text-Processing[17]

It is a HTTP web service for sentiment analysis, Part-Of-Speech (POS) Tagging and chunking, Phrase Extraction and Named Entity Recognition. It is free and available for non-commercial use without any authentication. It classifies sentences into three classes' positive negative and neutral. It provides with probability of pos and neg which sums up to 1 and neutral has been given standalone score. This supports three languages English, Dutch and French. It takes maximum of 80,000 characters. This service uses NLTK for sentiment analysis. Here's a glimpse (Fig.7).

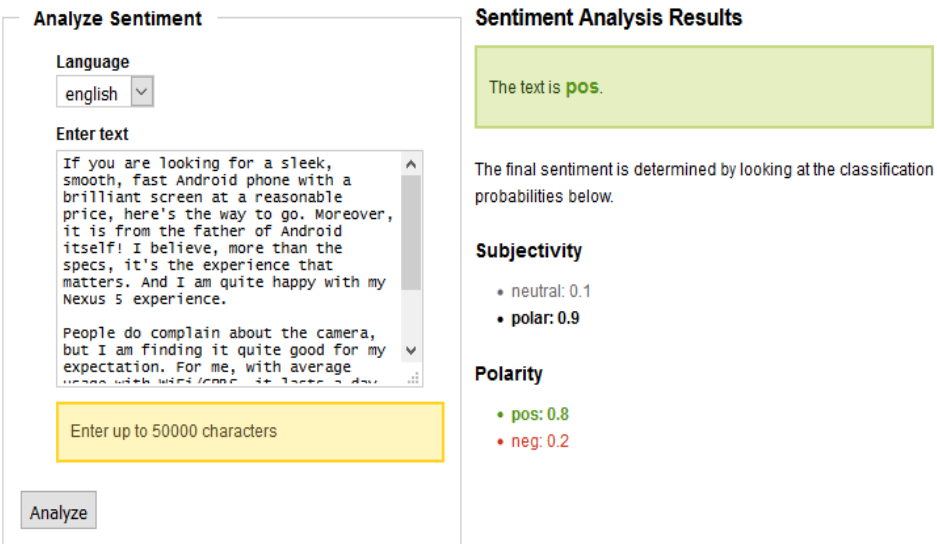

Fig.7: Sample amazon review and its output in Text-Processing

A comparative analysis has been tabulated in Table 1 below.

Tools	Document-Level	Sentence-Level	Word-Level	Emotion Classification	Length constraints	Opinion Level	Languages Supported
Opendover	Yes	Yes	No	Yes	No	[-9,9]	English
SentiStrength	No	Yes	Yes	Yes	Yes	[-5,5]	English, Finnish, German, Dutch, Spanish, Italian and Russian
Lexalytics	Yes	Yes	Yes	No	No	[-1,1]	English, French, Arabic, German, Hebrew, Japanese, Korean, Italian
Sentimetrix	No	Yes	Yes	No	Yes	[-1,1]	English, French, Italian, Chinese, Russian, Spanish
Text-Processing	No	Yes	Yes	No	Yes	[pos, neg]	English, Dutch and French

Table 1: A comparative analysis of various features of different opinion mining web services

4. Experiments

4.1 Data Collection

The above-mentioned web services perform multiple tasks. The tasks include POS tagging, sentiment analysis, entity recognition and sentiment rating prediction. In this work, we are comparing such web services in respect of sentiment classification into positive, negative or neutral and their intensities applied on amazon customer reviews. Reviews are multi-domain product reviews ranging from mobile phones, furniture, clothes, shoes and accessories. Each review is highly polar provided with star rating between [1,5]. Those reviews are taken which had complete sentences and provide strong opinions regarding products.

1000 such reviews had been taken out of which 700 are used for training and 300 are used for testing. Data can be taken large but there's a constraint on these web services that provide free results for specific number of documents or

sentences. Each service takes time between two consecutive analysis and say specific number of calls/day. So, documents are selected on equal amounts for negative and positive each day. Reviews contains 465 positive reviews, 82 neutral and 453 negative reviews. Their characteristics are shown in Table 2.

Polarity	Sentences	Words	Avg. Positive Sentences in each Review	Avg. Negative Sentences in each Review
Positive	2,118	33,468	2.8	1.3
Negative	1,504	10,280	3.6	0.8
Neutral	142	1,232	0.4	0.3

Table 2: Detailed description of review dataset

4.2 Evaluation Metrices

To evaluate these reviews, we have taken these measures:

1. *Accuracy*

$$accuracy = \frac{\#reviews\_classified\_correctly}{\#total\_reviews}$$

Where #reviews_classified_correctly represents the number of reviews either positive, negative or neutral are classified correctly in this experiment and #total_reviews is the total reviews taken.

2. *Precision*

It is a measure that can be calculated separately for each class. So, Precision is calculated for each class. e.g. positive precision can be calculated as:

$$Precision_{pos} = \frac{\#pos\_reviews\_classified\_correctly}{\#total\_reviews\_classified\_positive}$$

Where #total_reviews_classified_positive represents the number of negative, positive and neutral reviews classified as positive. Similarly, we can find precision for negative and neutral reviews.

3. *Recall*

This measure is also calculated separately for each class. This has been computed considering only one class at a time.

$$Recall_{pos} = \frac{\#pos\_reviews\_classified\_correctly}{\#total\_positive\_reviews}$$

Where #total_positive_reviews are total number of positive reviews submitted in the work for results. Recall for negative and neutral are computed similarly.

4.3 Experiment Setup

Reviews and sentences are submitted to each tool in their default configuration as each tool allow only specific settings as freeware, so only some parameters are in control and in some case none. Some of the reviews were not analysed fully because of some constraints in each tool. Constraints like length of reviews or sentence that can be submitted at once, calls/day, some characters are not recognized by some of these tools. Such reviews are not considered for evaluation purpose. e.g. SentiStrength only takes a small sentence at once. So, it takes various turns to evaluate each review. For this tool, each sentence is evaluated and review has been labelled as positive or negative depending upon maximum number of sentences labelled as positive or negative.

The results so obtained from each tool is either a character string as {positive, negative, neutral} or numeric value ranging from [-5, 5], [1, 10] or [0, 1], thus, each value is normalized into the range [-1,1].

4.4 Results

Results are commented only based upon polarity classification and sentiment classification ignoring all other features that these tools support. Results have been tabulated in Table 2. Important information is conveyed by Neutral class as maximum of the tools are unable to classify these reviews correctly. Major of the reviews generally contains some kind of polar words due to which tools classify them as either positive or negative. Results are tabulated in Table 3.

Service	Accuracy(%)	$Precision_{pos}$ (%)	$Precision_{neg}$ (%)	$Precision_{neutral}$ (%)	$Recall_{pos}$ (%)	$Recall_{neg}$ (%)	$Recall_{neutral}$ (%)
Opendover	54.6	42.2	62.1	3.7	47.8	58.2	13
SentiStrength	61.8	54.5	**78.2**	1.2	59.3	76.1	2.5
Lexalytics	**64.7**	**68.9**	72.8	**12.3**	67.1	68.2	**15.4**
Sentimetrix	51.3	42.9	74.3	0	**69.1**	51.5	0
Text-Processing	63.2	56.7	68.9	2.3	59.4	**78.9**	0.6

Table 3: Results obtained from different web tools

5. Conclusion

The given work provides a detailed analysis of some free web services that provide sentiment analysis. Most of the tools put restrictions on free access to their functionalities and number of times they can be accessed. So, these tools are beneficial for research purpose or for those who don't want to create altogether a new algorithm for sentiment analysis and wish to use already created APIs in their work. As the above mentioned web services provide their APIs and free access to their online service, it is quite easy to use them.
Also, various services provided by each web-service is enlisted here, so, one can choose between various tools according to their need. If one wants to use part-of-speech with sentiment analysis then it one may choose Text-Processing or other. Therefore, these services have been analysed and assessed under different conditions. To fulfil this work, data from amazon reviews have been collected which is a mix of highly opinionated and some neutral reviews. Their tagging has been done using star rating and already provided distinguished reviews.
The above results show that Lexalytics and Text-Processing have provided better results under different conditions and therefore, can be used with almost every review. SentiStrength has a major drawback that only a short sentence can be processed using web service, so it has affected its accuracy. But almost all the tools provided better results for negative reviews in comparison to positive and neutral reviews.
Major drawback of these tools is being their inefficiency in detecting neutral reviews. Most of these tools suffer from this problem due to which sentences which are bipolar are classified as neutral. Lexalytics, Opendover and SentiStrength are efficient in providing the sentiments about a particular entity.
Therefore, this work demonstrate that users can use these web services purposefully for their work and will be able to get satisfactory results. Also, these tools have lot of challenges ahead. They suffer from many problems like absence of degree of polarity, amount of characters that can be entered or some tools simply calculating the polar words without any relationships.

Acknowledgements

This work was supported in part by the Council of Scientific and Industrial Research (CSIR), New Delhi with file no. 09/263(1049)/2015-EMR-I.

References

1. Resnick P, Varian HR, Recommender Systems. Communications of the ACM vol. 40(3), 1997. Pp. 56-58.
2. Balahur A., Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications, Computer Speech and Language. Vol. 28(1), 2014. Pp. 1-6.
3. Rush AM, Chopra S, Weston J. A Neural Attention Model for Abstractive Sentence Summarization. Proceedings of EMNLP 2015.
4. Pang, B. Lee L.*Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval. Vol 2(1-2), 2008. 1-135
5. Liu, B. (2010). Sentiment Analysis and Subjectivity. *Handbook of Natural Language Processing*, (1), 1–38. http://doi.org/10.1145/1772690.1772756
6. https://en.wikipedia.org/wiki/Opinion
7. Yang, S., & Ko, Y. (2011). Finding relevant features for Korean comparative sentence extraction. *Pattern Recognition Letters*. http://doi.org/10.1016/j.patrec.2010.09.003
8. Whitelaw, C., Garg, N., & Argamon, S. (2005). Using appraisal taxonomies for sentiment analysis. *Proceedings of CIKM*, 625–631. http://doi.org/10.1145/1099554.1099714
9. Gaspar, R., Pedro, C., Panagiotopoulos, P., & Seibt, B. (2016). Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Computers in Human Behavior*, *56*, 179–191. http://doi.org/10.1016/j.chb.2015.11.040.
10. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive Language Detection in Online User Content. In *WWW '16 Proceedings of the 25th International Conference on World Wide Web* (pp. 145–153). Canada. http://doi.org/10.1145/2872427.2883062.
11. Lexalytics. Available from: https://www.lexalytics.com/ accessed on
12. Opendover. Available from: http://opendover.nl/ accessed on
13. Sentistrength. Available from: http://sentistrength.wlv.ac.uk/ accessed on
14. Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544–2558.
15. Sentimetrix. Available from: http://www.sentimetrix.com/ accessed on
16. Gate. Available from: https://cloud.gate.ac.uk/ accessed on
17. Text-Processing. Available from: http://text-processing.com/ accessed on