JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SENTIMENTAL ANLYSIS ASPECTS WITH ITS APPLICATIONS AND CHALLENGES

Dr. Kusuma Kumari B.M

Assistant Professor, University College of Science, Tumkur University
Tumakuru, Karnataka- 572103,India
kusuma.kuku@gmil.com

Prof. Ramakanth Kumar .P

Professor, Department of Computer Science and Engineering, R.V. College of Engineering, Bangalore, Karnataka, India
ramakanthkp@rvce.edu.in

Abstract
People are becoming tech savvy these days. Most of the people begin their day by posting opinion about various entities on social media, sites etc. This has resulted in explosion of abundant unstructured data which has attracted researchers to extract knowledge out of this data to gain insight. Sentiment analysis is performed on data gathered from twitter, face book, review sites etc to understand the feelings of people. In this paper we discuss the procedure to carryout sentiment analysis, different levels of sentiment analysis, applications and its challenges.

***Keywords*: Sentiment analysis; Sentiment analysis; social media; applications; challenges.**

1. Introduction

The 21st century enables data to flow across all sectors. This has allowed the birth of BigData[8]. In recent years the number of people using the internet has increased exponentially. People not only browse the data available over internet but they tend to produce enormous amount of data. Social media has become the inseparable part of most of the people's life. Every minute huge amount of data is been produced over social media. This data can be used in a productive way to analyse various aspects. Twitter is one of the most popular social network website where people are free to post their views related their subject of interest. The huge number of daily posts on social media has paved the way to many research areas such as sentiment analysis.

2. Sentiment Analysis

The Sentiment Analysis is a process to identify the opinion of people towards specific product, persons, services, organizations, news, events, issues etc. It is also called as subjectivity analysis, opinion mining, and appraisal extraction. Through sentiment analysis one can determine whether a piece of text has positive, negative or neutral opinion about something
Sentimental analysis is a procedure to find the polarity of the text and the significant classified into three types: positive, negative, and neutral. Sentimental analysis requires the use of performance in the training group and the exact value of the specific text. A lot of research on the part of sentimental analysis mainly works on social media sites and marketing levels[4].

Everyday enormous amount of data is being generated by people and it's almost impossible for the human to analyze it. Therefore this unstructured data is automatically analyzed using various sentiment analysis techniques.
Over the years number of researchers has worked in this field. Sentiment analysis is performed by using techniques like Natural Language Processing (NLP), Machine Learning, Text Mining and Information Theory and Coding, Semantic Approach.

The opinions posted by users on different social media and other platforms can be classified into positive, negative or neutral by using different techniques of sentiment analysis that are mentioned above. Normally these kinds of data are unstructured data therefore needs pre-processing [6] [1].

3. Steps to Analyze Sentiment Data

Since sentiment analysis is a very complicated process to analyse huge voluminous data yet it can grouped under five different steps. They are:

Data collection: The first step involves gathering data from different sources. The data required for analysis can be collected from different sources such as blogs, forums, social network sites, app stores. The data collected from these sources are highly unstructured and it becomes very difficult to analyse manually therefore natural language processing techniques are used [5].

Text preparation: The second step involves cleaning the text. It includes removal of irrelevant data that don't add sentiment value.
Sentiment Detection: In the third step the reviews are thoroughly examined to identify subjective and objective sentences.

Sentiment Classification: In this step subjective sentences are identified into positive, negative or neutral .
Presentation of output: The main goal of performing sentiment analysis is to make sense out of massive amount of data. The final step in sentiment analysis to show the result of analysis through different graphs such as histogram, line graphs or bar graphs etc.

4. Levels of Sentiment Analysis

The sentiment analysis can be done in various levels such as document level, sentence level, and aspect/feature level. These levels are discussed in detail in the following sections.

4.1. *Document level sentiment analysis*

In document level sentiment analysis entire review is considered to extract sentiment and the entire opinion is categorized based on overall sentiment of the opinion holder. If the analyser wants to extract opinion on a single entity then the document level analysis is best suited since single person writes a document. In this type the whole document is considered to perform sentiment analysis. Initially preprocessing is performed to remove unnecessary sentences. Both supervised and unsupervised machine learning classification methods are used for carrying out document level sentiment analysis. To train the system supervised machine learning algorithm such as Support Vector Machine (SVM), Naive Baye's, KNN and Maximum Entropy can be used. The reviewer rating is collected in the form of one to five stars and review text can be used to train and test dataset.

The various features that can be used for the machine learning are term frequency, document frequency, Part of speech tagging, Opinion words, opinion phrases, negations and dependencies. To carryout unsupervised machine learning opinion words inside the document are extracted. The semantics of the extracted words can be found with the help of point-wise mutual information [3][9][7].

4.2. *Sentence level sentiment analysis*

The sentence level sentiment analysis usually consists of two steps:

- The sentence is classified based on subjectivity classification such as objective and subjective.
- The subjective sentence is further classified into sentiment such as positive and negative.

The objective sentences are those that contain factual information whereas subjective sentences consist of feelings of opinion holder. One of the methods used for subjective sentence classification is Naive Bays classification method.

4.3. *Aspect/Feature level Classification*

In Aspect/Feature Level Classification the objective features of objective sentences are identified and extracted.

5. Source of Data

The quality of products and services provided by various entities are well understood by analyzing the user's opinion posted on different review sites, blogs etc.

5.1. *Review sites*

The users post reviews about products, services, people and businesses on review sites. These reviews are extracted to perform sentiment analysis. This helps to improve the business [3].

5.2. *Blogs*

The blog is an online informational website where users post their opinions. Usually recent information is displayed on top of the webpage. There are many different kinds of blogs such as personal blogs, collaborative blogs, micro blogging, corporate blogs etc., contains different kinds of contents.

5.3. *Forums*

The forum is an online discussion site where people converse with each other by posting messages. Generally forums are restricted to single topic which is helpful in doing sentiment analysis on a single domain.

5.4. *Social Networks*

Social networking sites are online platforms where people share their opinions of interest. People share their feelings on social networking sites such as twitter, facebook etc resulting in production of bulk of data and identifying sentiment from it becomes real challenge.

5.5. *App Store*

Nowadays apps are designed for specific operating systems such as Android, iOS, macOS, windows. App stores are online stores where users can browse through various apps, view information about app and acquire the app. App stores allows the users to give review and rating for apps. These reviews and ratings are used by other users, developers and app store owners. The app store owners analyse the reviews to detect apps that are not good and malicious developers.

6. Applications of Sentiment Analysis

In today's virtual world people are free to give feedback on any product or service they are interested in, resulting in voluminous data which is impossible for humans to analyse manually. Hence sentiment analysis has got wide range of applications in various fields.

6.1. *Use of online reviews*

Online reviews are very useful for the customers to purchase new products by considering the feedbacks of other customers. It is also helpful for vendors to get the overall feedback for a particular product to improve their service for the customers [2].

6.2. *Used for recommendation*

The items previously purchased or selected or ratings given to the items are referred to as past behavior of customers. These past behavior are collected to analyze what customer's likes and dislike in order recommending them the relevant products or services.

6.3. *Used in business intelligence*

The data collected about the opinions about various products or services are analysed gain insights into business process. These insights are used make strategic plans to increase the profitability.

7. Challenges in Sentiment Analysis

The data produced through social network sites, blogs, forum, app stores are highly unstructured and hence it is a very challenging task for the analyser to make sense out of this data.

- The unstructured data contains lot of noise due to the usage of misspellings and abbreviations like "lol" (laugh out loud),"RIP"(Rest In Peace),"tysm"(thank you so much) etc.
- Mixed opinions can be difficult to classify for example "The performance of the mobile is very good but the battery heats up quickly".
- People sometimes make use of sarcasm in review to criticize the product for example "The speaker is so good that I can hear sounds from other galaxies! Totally worth the money". Detection of sarcasm in language is a challenging problem.
- When same event is discussed in different data sources like Twitter, Face book and YouTube, if semantic relationship is constructed out of these data then it can be of great help in accomplishing insight and better understanding of the event[10].

8. Conclusion

The data produced on social media are multiplying day by day and extracting knowledge out of this data has got various applications that lead to the way for sentiment analysis. The sentiment analysis is a complicated process which involves lot of procedures to extract relevant facts from big data. The data can be collected from different sources that are mostly unstructured and involves many challenges to analyse it.

References

[1] Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. International Journal of Computer Applications, 125(3).

[2] Kharde, V., & Sonawane, P. (2016). Sentiment analysis of twitter data: a survey of techniques. arXiv preprint arXiv:1601.06971.

[3] Moralwar, S., & Deshmukh, S. (2015). Different Approaches of Sentiment Analysis. International Journal of Computer Sciences and Engineering, 3(3), 160-165.

[4] Nitesh Sharma, Ussama Yaqub, Vijayalakshmi Atluri, Soon Ae Chun, Rachit Pabreja, and Jaideep Vaidya " Web-based Application for Sentiment Analysis of live Tweets" ACM, DOI: htttps://doi.org/10.1145/3209281.3209402, Year: 2018

[5] Nitin Indurkhya, Fred J. Damerau, "Handbook of Natural Language Processing", Second Edition, CRC Press, 2010

[6] Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In LREc (Vol. 10, No. 2010, pp. 1320-1326).

[7] Patni, S., & Wadhe, A. (2014). Review Paper on Sentiment Analysis is–Big Challenge. International Journal of Advance Research in Computer Science and Management Studies, 2(2).

[8] R. Addo-tenkorang and P. T. Helo, ''Big data applications in operations/supply-chain management: A literature review,'' Comput. Ind. Eng., vol. 101, pp. 528–543, Nov. 2016.

[9] Raisa Varghese1, Jayasree M2, "A SURVEY ON SENTIMENT ANALYSIS AND OPINION MINING", IJRET:International Journal of Research in Engineering and Technology ISSN: 2319-1163 | ISSN: 2321-7308.

[10] Tsur, O., Davidov, D., & Rappoport, A. (2010, May). ICWSM—a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In fourth international AAAI conference on weblogs and social media.