A Review: Sentiment Analysis and Opinion Mining

Sakshi kalra	**Rajesh Sachdeva**	**Anjali Dhawan**
Assistant Professor	**Assistant Professor**	**Assistant Professor**
Dev Samaj College For women	**Dev Samaj College For women**	**Dev Samaj College For women**
Ferozepur	**Ferozepur**	**Ferozepur**

Abstract: Web has provided a platform where users are free to give their opinions, suggestions, and remarks openly. With the help of web services now customers can express their experience about E-shopping, E-payment, Likes or dislikes related product or its services. This is the straightforward idea to deliver their observations. From these opinions system gathered the huge amount of data called actual sentiments .There are different stages are involved in sentiment analysis like collecting data from different resources, classification, combining or grouping together and then evaluate realistic values. It is completely automatic process in which system extract knowledge from user's opinion regarding any company, specific product or its features. This paper mentioned all the details regarding sentiment analysis, collection of data from different resources, current utilization in different fields and the working process and also emblematizes the hurdles which are being faced by the opinion mining.

*Keywords: web mining, data sources, working process, current utilizations, challenge*s

1 Introduction

Use of internet is amplifying today. This automated system helps each and every viewer to generate such a vast amount of data regarding any field like science, medical, engineering, marketing finance etc. With the help of theses services users are free to explore their opinions. Customer responses on the sites whether positive or negative can have the power to change the other customer's decisions, as well. These responses have become vital information for business and also help to take decisions for better marketing and product development plans.

Sentiment analysis is used to find the clues that present the true picture of customer's mind, this picture shows what type of product or services they actual want [4] through blog posts, comments, reviews or tweets. This analysis process can be useful in many ways like it helps to deciding which product is most favorable and which feature is like by countless persons and even find how many no of persons which are not fully satisfied regarding company, product or its service.

2 Data Source

User opinion is a most important factor for the improvement of the quality of a product or Services. For this we have blogs, review sites, data and micro blogs that provide a better perceptive way of the products and services provided to customers.

2.1 Blogs

Blog pages/blogging is rising rapidly. They are becoming the most popular means to express opinions/ personal opinions regarding business, politics.

2.2 Review sites

When user purchases something, the reviews/ratings of other users can play an important role in purchasing decisions. User also wants to check the actual reviews on the internet of a particular product for make sure that he or she is on the right track. The reviewer's data mostly collected from commercial sites. [4]

2.3 Data Set

Positive, negative or neutral opinions are contained in dataset extracted from e-commerce sites, including books, DVDs. for classifying the movie reviews and many more are available as dataset.

2.4 Micro-blogging

Twitter is a popular micro-blogging service where users share their ideas or views called "tweets". These tweets express opinions about different topics. These messages are also used as a source for classifying user's sentiment for eg. Famous personalities like movie stars, politicians are connecting their selves with public to express their emotions via these web services.

3 Work flow of Sentiment analysis
3.1. Extraction
Extraction means data collection from different sources. In this phase data is collected from social sites like Twitter where number of users share their opinions called tweets. These tweets are the sentiments or opinions of user's. The main theme of this process is basically to analyzing the data and converting it into knowledge. In this process firstly collect the data from social sites, check the message length and also check cyber slang (colloquial speech) in these messages after checking all the details these messages are become a real time representation of the sentiments.
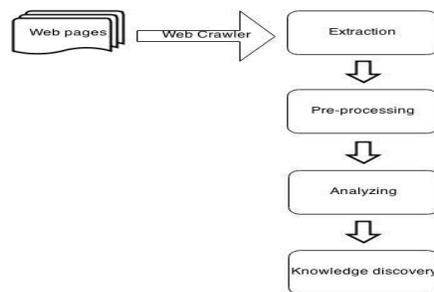

Fig.1 Working of sentiment analysis

3.2. Pre-Processing
Before analyzing the next step is to clean the extracted data. All the errors are checked and removed in this phase to make sure that essential data is preserved and rest is removed. Some techniques are used for pre-processing.

3.2.1 Supplanting emoticon
Due to the restriction of words, users use sticker's smileys and different types of icons rather than text to depict their emotions.

3.2.2 Uppercase and Lowercase Identification
Strong emotions and feelings like anger, love, and affection are always expressed in uppercase and shows the intensity of the emotions. E-shouting is used as indicator to decide the polarity of the text. We also observe inconsistent casing (e.g. HeLLoO) in texts on social media. It is necessary to make sure that there is consistency in casing of the texts. [1]

3.2.3 Extraction of URL
Due to the restriction that limited data is to be send that's why user use URL in order to share something extra .The information obtained from the URL basically supports the sentiments which are expressed in tweets. But it consumes more time and also costly. Therefore a compact equivalent class <URL> can be used.

3.2.4 Punctuations:
The basic purpose of the user who is sharing something is not bother about sentence formation but only give stress on emotions expressed in the posts. E.g. Hurrah! - Oh! Wow! , to express strong feeling. In this technique of pre-processing, punctuation marks are removed.

3.2.5 Compression of words:

To describe the emotions clearly, users stretch the words. For example- "food is sooooooo delicious".It gives focus of that thing which is delicious. Therefore, we can reduce it to shorter sequence like „soo".

3.3. Analysis

After pre-processing stage, next step is sentiment analysis. System found the sentiments contained in the data. The words or keywords describing emotions are used to be valuate at this stage. During the analysis if an existing word is seen repeated in a sentence, it is record and is added to the value of the entire text to decide the polarity of the text which makes better understanding about the emotions. Most of the algorithms for sentiment analysis use basic terms to review and express opinions about a service or product .So at this stage, many more factors are taken into consideration to get a precise result about the true opinion whether it is positive or negative.

3.4. Knowledge Discovery

At this stage, system finds the true opinion of the people, it is essential to store the data which is related to the affair. Once the data is collected it can be used to generate charts or graphs. The knowledge gathered from these charts or texts collected from the web would help to formulate market strategies.
After all these stages are completed, the process of sentiment mining is successfully executed.

4. Current Utilization

The applications assist in making sense of hundreds of applications covering several fields are contained in sentiment mining. [5]Sentiment mining works diversely from the conventional survey methods and depends on listening in spite of asking which depicts more accurate results. Sentiment Analysis helps corporate sector to get customer opinion in real-time. It can be used in different fields for various purposes.

4.1 Online Commerce

Sentiment analysis is commonly used in e-commerce. Websites allows their users to submit their experience about shopping and product qualities. They provide summary of the product with ratings or scores. Customers can easily view opinions and recommendation information as well as specific features of products. Product images are also display for users.

4.2 Voice of the Market (VOM)

Voice of the Market determines the feeling of the customers about products or services .Appropriate information from the market helps in gaining advantage and new product development. Timely information is help to reach or target the market goals, design new marketing strategies, improve product features .It can also helps to anticipate if any chances of product failure. According to Zhang et al. [6] weakness finder system are available, which can help manufacturers find their product weakness from Chinese reviews by using aspects based sentiment analysis. There are some of the sentiment analysis services are available like Radiant6, Sysomos, Viral heat, Lexalytics etc and some free tools are also available.
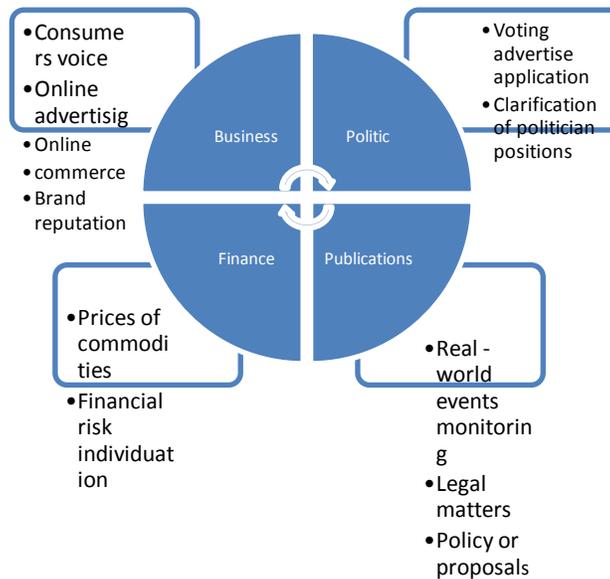
Fig.2 Applications of sentiment analysis

4.3 Voice of the Customer (VOC)

Voice of the Customer means analyzing the reviews and feedback of the customers. It is a main element of Customer Experience Management. Which helps in identifying new opportunities for product inventions. Customer's opinions also help to identify functional and non-functional requirements of the products like performance and cost.

4.4 Brand Reputation Management

Brand Reputation Management is concerned about managing a reputation in market. Opinions from customers or editors can damage or enhance the market reputation. The main aim of reputation management is to figure out the public observation regarding companies and its products. In other words brand reputation refers to how a particular brand product is viewed by others. A trusted brand reputation means consumers trust your company, and feel good about purchasing your goods or services and an unfavorable brand reputation, however, it may cause users to not trust your company and be hesitant about purchasing your products or services. While companies used to be able to improve their brand reputation with ease using traditional advertising and public relations. Sentiment analysis helps in determining how company's brand, product or service is being perceived by community online.

4.5 Government

When a government wants to assess their strength and weaknesses it is done through sentiment analysis. For example, "If this is the state, how do you expect truth to come out? The officer who is investigating scam himself is deeply corrupt." this example shows negative sentiment about government. Whether it is tracking citizens' opinions on a new 108 system, identifying strengths and weaknesses in a recruitment campaign in government job, assessing success of electronic submission of tax returns, or many other areas, we can see the potential for sentiment analysis.

5   Hurdles in Opinion Mining

Opinion mining has become really important as nowadays for devising even smaller strategies, corpuses of data has to be analyzed. The process of opining mining is a great tool in it. Sentiment mining reducing the burden of human shoulders. There are many techniques available for sentiment analysis. But each

technique has its own issues and certain challenges. Mainly there are two techniques used for opinion mining. [3]

1) Lexicon based and 2) Learning based

High precision is involved in Lexicon based technique but gives a low recall. Secondly, it is not available in all languages. Labeled examples used to classify text is one of the step of Learning based techniques. But it needs a great deal of training and its dataset. Yet another one is Syntactic technique which yield god results no language independence is there. Other challenges that erupt during the process of classification:-

1. Product reviews, comments and feedback could be in different languages (English, Urdu, Arabic, and French etc) therefore to handle each language according to its orientation is a challenging task.

2. As noun words are considered as feature words but Verbs and adjectives can also be used as feature words which are difficult to identify.

3. If a customer-One comments that, "the Picture quality is excellent" and customer-Two comments, "Picture quality of LED is very good". Both are talking about same feature but with different wording. To group the synonym words is also a challenging task.

6. Opinion words could be different according to situation. For example "Camera size of mobile phone is small". Here adjective small used in positive sense but if customer parallel said that "the battery time is also small". Here small is represent as negative opinion for battery backup. To identify the difference between same adjective words in different situation is also a challenging task. [5]

7. Generally Customer use short words in their conversations or sometimes in reviews. For example: - u for you, *pic for picture, f9 for fine, b4, before, gud for good etc.* To deal with such type of language need a lot of work to mine opinion.

8. Different people have different writing styles; same sentence may contain positive as well as negative opinion, so it is difficult to parse such type of sentences.

9. Spam messages or fake reviews have become really common and used on a large scale on social websites. This creates a hurdle in the process of sentiment mining. So one of the biggest challenges is to identify such spam messages and fake reviews .[3]

10. Filter bubble, Skewness and Asymmetry in opinion mining are another challenges in Sentiment mining.

6. Conclusion:

In a nutshell we can say that Sentiment detection has a wide variety of applications in information systems that includes real time applications, reviews based on classifications and summarization. It has varied diversity of applications that could prove to be advantageous to many fields such as marketing, business analytics, and knowledge bases and so on. The major challenges that may exist in use of other languages are dealing with negation expressions; produce a summary of opinions that are based on features/attributes of a product, complexity of sentence/ document, handling of implicit product features, etc. As it involves natural language processing, Sentiment mining arises as a challenging field with many hindrances. In order to overcome their individual drawbacks and benefit from each other's merits, and in order to enhance the sentiment classification performance there are many research work in progress and soon will be unveiled in coming years to overcome the current problems and enable smooth functioning eliminating the challenges.

References:

,

1. Procedure of Opinion Mining and Sentiment Analysis: A Study Rushabh Shah Á*and Bhoomit PatelÀÁInformation Technology,Dwarkadas J.Sanghvi College of Engineering,Vile Parle(W),Mumbai-400056, IndiaAccepted 02Dec2014, vailable online 10Dec2014, Vol.4, No.6 (Dec2014)

2. An Analysis on Opinion Mining: Techniques and Tools G.Angulakshmi 1, Dr.R.ManickaChezian2Research Scholar, Department of Computer Science, NallamuthuGounderMahalingam College, Pollachi, India1Associate Professor,Department of Computer Science, NallamuthuGounderMahalingam College, Pollachi, India

3. Opinion Mining , Analysis and its Challenges Nidhi R. Sharma , Prof. Vidya D. Chitre Computer Department, University of Mumbai,Computer Department, University of Mumbai

4. Volume 2, Issue 6,June2012ISSN: 2277 128XInternational Journal of Advanced Research inComputer Science and Software EngineeringResearch PaperAvailable online at: www.ijarcsse.comSentiment Analysis and Opinion Mining:A SurveyG.Vinodhini*Assistant Professor, Department of Computer Science and Engineering, Annamalai University,.ChandrasekaranProfessor, Department of Computer Science and Engineering, Annamalai University,

5. Research Challenge on Opinion Mining and Sentiment Analysis*David Osimo1and Francesco Mureddu2

6. International Journal of Engineering Research & Technology (IJERT)Vol. 2 Issue 2, February-2013ISSN: 2278- 0181Applications and Challenges for Sentiment Analysis : A Survey1 Mr. Saifee Vohra, 2 Prof. Jay Teraiya

7. *International Journal of Computer Applications (0975 – 8887) Volume 125 – No.3, September 2015* 26 Approaches, Tools and Applications for Sentiment Analysis Implementation

8. Raisa Varghese, Jayasree, "A Survey on Sentiment Analysis and Opinion Mining", International Journal ofResearch in Engineering and Technology (IJRET), Vol 2 Issue 11 Nov 2013.[8]Sindhu, Chandrakala, "A Survey on Opinion Mining and Sentiment Polarity Classification", International Journal of Emerging Technology and Advanced Engineering.Vol 3 Issue 1,Jan 2013

9. Walaa Medhat , Ahmed Hassan , Hoda Korashy (2014)Sentiment analysis algorithms and applications:A survey . Ain Shams Engineering Journal

10. B. Seerat and F.Azam "Opinion Mining: Issues and Challenges (A survey)" International Journal of Computer Applications (0975 –8887) Volume 49–No.9, July 2012.