International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 07, July 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

SENTIMENT ANALYSIS: AN ANALYSIS ON ITS PAST, PRESENT AND FUTURE SCOPE

Sanjeev Kumar Sharma

Department of computer science and applications, DAV University,Jalandhar, (India)

ABSTRACT

Opinion mining or Sentiment analysis is used to identify user's view. It is a process of judgment and evaluation to extract subjective information from a text file or a text document. The important task of opinion mining is to extract the emotions according to the polarity and neutrality of text. Social networking websites like Twitter, LinkedIn and many more provide the facility to user to share their thoughts via text messages i.e. updating post and status. Sentiment analysis or Opinion mining plays an important role in finding the area of interest based on user's previous actions.in this paper author has given the present past and future scope of the sentiment analysis. The research scope of opinion mining is also discussed.

Keywords: Opinion Mining, Sentiment Analysis, social networking.

I. INTRODUCTION

Social networking portals have been widely used for expressing opinions in the public domain.Sentiment analysis is one of the key emerging technologies in the effort to help people navigate the huge amount of user generated content available online. It is true that via these media citizens can express their desires, problems, emotions and feelings and the experts can make use of it by properly extracting and analyzing it. But the extraction and analysis of huge unstructured internet content is beyond the human power and time. The content is mostly written in natural language. This situation necessitate an automatic natural language processing tool that extract and analyze the people sentiments from this unstructured texts. Numerous researches are undergoing in this direction. This research domain is called Opinion mining and sentiment analysis.

II. OPINION MINING AND SENTIMENT ANALYSIS

It is an extension of data mining which utilizes natural language processing techniques to extract people's opinion from World Wide Web. The recent trend in internet that encourages users to contribute their opinion and suggestion created a huge collection of valuable information in the web. The Opinion mining system analyse each text and see which part contain opinionated word, which is being opinionated and who has written the opinion.Sentiment analysis analyses each opinionated word or phrase and determines its sentiment polarity orientation, whether it is positive or negative or neutral. It gives the summarized opinion of a writer or speaker. Sentient analysis can be done at word level, sentence level and document level

III. BASIC TERMINOLOGY OF SENTIMENT ANALYSIS

Formally stating Sentiment Analysis is the computational study of opinions, sentiments and emotions expressed in text. The goal of sentiment analysis is to detect subjective information contained in various sources and

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

determine the mind-set of an author towards an issue or the overall disposition of a document. It is described subjectivity as the linguistic expression of somebody‘s opinions, sentiments, emotions, evaluations, beliefs and speculations. The words opinion, sentiment, view and belief are used interchangeably but there are subtle differences between them.

- *Opinion:* A conclusion thought out yet open to dispute (each expert seemed to have a different opinion).
- *View:* subjective opinion (very assertive in stating his views).
- *Belief:* deliberate acceptance and intellectual assent (a firm belief in her party‘s platform).
- *Sentiment:* a settled opinion reflective of one‘s feelings (her feminist sentiments are well-known).

Sentiment analysis is done on user generated content on the Web which contains opinions, sentiments or views. An opinionated document can be a product review, a forum post, a blog or a tweet, that evaluates an object. The opinions indicated can be about anything or anybody, for e.g. products, issues, people, organizations or a service. The same can be mathematically represented an opinion as a quintuple (o, f, so, h, t), where o is an object; f is a feature of the object o; so is the orientation or polarity of the opinion on feature f of object o; h is an opinion holder; t is the time when the opinion is expressed.

Object: An entity which can be a product, person, event, organization, or topic. The object can have attributes, features or components associated with it. Further on the components can have subcomponents and attributes.

Feature: An attribute (or a part) of the object with respect to which evaluation is made.

Opinion orientation or polarity: The orientation of an opinion on a feature f indicates whether the opinion is positive, negative or neutral. Most work has been done on binary classification i.e. into positive or negative. But opinions can vary in intensity from very strong to weak. For example a positive sentiment can range from content to happy to ecstatic. Thus, strength of opinion can be scaled and depending on the application the number of levels can be decided.

Opinion holder: The holder of an opinion is the person or organization that expresses the opinion

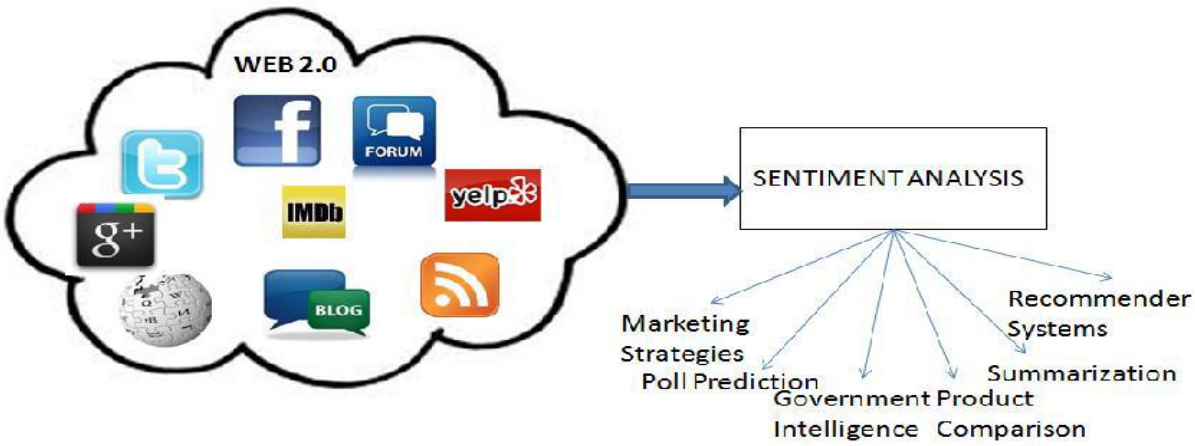

Fig: Classification of Sentiment Analysis

IV. BACKGROUND

The term Web 2.0 was made popular by Tim O‘Rielly in the O'Reilly Media Web 2.0 conference in late 2004. Web 2.0 is an evolution from passive viewing of information to interactive creation of user generated data by the collaboration of users on the Web. Every facet of Web 2.0 is driven by contribution and collaboration. The

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

evolution of Web from Web 1.0 to Web 2.0 was enabled by the rise of read/write platforms such as blogging, social networks, and free image and video sharing sites. These platforms have jointly allowed exceptionally effortless content creation and sharing by anyone. The research field of sentiment analysis has been rapidly progressing because of the rich and diverse data provided by Web 2.0 applications. Blogs, review sites, forums, microblogging sites, wikis and social networks have all provided different dimensions to the data used for sentiment analysis.

4.1 Review Sites

A review site is a website which allows users to post reviews which give a critical opinion about people, businesses, products, or services. Most sentiment analysis work has been done on movie and product review sites . The purpose of a review is to appraise a specific object, thus it is a single domain problem. Sentiment analysis on review sites is useful to both manufacturers and potential consumers of the product. The manufacturers can gauge the reception of a product based on the reviews. They can derive the features liked and disliked by the reviewers.

4.2. Blogs

The term web-log or blog, refers to a simple webpage consisting of brief paragraphs of opinion, information, personal diary entries, or links, called posts, arranged chronologically with the most recent first, in the style of an online journal. The bloggers post at hourly, daily or weekly basis which makes the interactions faster and more real-time. Different blogs have different styles of presentation, content material and writing techniques. Sentiment analysis on blogs has been used to predict movie sales, political mood and sales analysis.

4.3.Forums

Forums or message boards allow its members to hold conversations by posting on the site. Forums are generally dedicated to a topic and thus using forums as a database allows us to do sentiment analysis in a single domain.

4.4.Social Networks

Social networking is online services or sites which try to emulate social relationships amongst people who know each other or share a common interest. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks Social network posts can be about anything from the latest phone bought, movie watched, political issues or the individual‘s state of mind. Thus posts give us a richer and more varied resource of opinions and sentiments.

4.4.1.Twitter

Twitter is an online social networking and micro blogging service that enables its users to send and read text-based posts of up to 140 characters, known as "tweets". Sentiment analysis on twitter is an upcoming trend with it being used to predict poll results among various other applications.

4.4.2. Facebook

Facebook is a social networking service and website launched in February 2004. The site allows users to create profiles for themselves, upload photographs and videos. Users can view the profiles of other users who are added as their friends and exchange text messages. Social media is the new source of information on the Web. It connects the entire world and thus people can much more easily influence each other. The remarkable increase

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

in the magnitude of information available calls for an automated approach to respond to shifts in sentiment and rising trends.

V. SENTIMENT ANALYSIS TASKS

Sentiment analysis is a challenging interdisciplinary task which includes natural language processing, web mining and machine learning. It is a complex task and encompasses several separate tasks, viz:

- Subjectivity Classification
- Sentiment Classification
- Complimentary Tasks
  - Object Holder Extraction
  - Object/ Feature Extraction

VI. APPLICATIONS AREAS OF OPINION MINING AND SENTIMENT ANALYSIS

Since the Opinion based or feedback based application are more fashionable, now a days, the natural language processing community shows much interest in Sentiment Analysis and Opinion Mining system. The explosion of internet has changed the people's life style, now they are more expressive on their views and opinions and this tendency helped the researchers in getting user-generated content easily.

The major applications of Opinion mining and sentiment analysis are the following:

- Purchasing Product or Service: While purchasing a product or service, taking right decision is no longer a difficult task. By this technique, people can easily evaluate other's opinion and experience about any product or service and also he can easily compare the competing brands. Now people don't want to rely on external consultant. The Opinion mining and sentiment analysis extract people opinion form the huge collection of unstructured content, the internet, and analyze it and then present to them in highly structured and understandable manner.
- Quality Improvement in Product or service: By Opinion mining and sentiment analysis the manufactures can collect the critic's opinion as well as the favorable opinion about their product or service and thereby they can improve the quality of their product or service. They can make use of online product reviews from websites such as Amazon etc..
- Marketing research: The result of sentiment analysis techniques can be utilized in marketing research. By sentiment analysis techniques, the recent trend of consumers about some product or services can be analyzed. Similarly the recent attitude of general public towards some new government policy can also be easily analyzed. These all result can be contributed to collective intelligent research
- Recommendation Systems: By classifying the people's opinion into positive and negative, the system can say which one should get recommended and which one should not get recommended.
- Detection of "flame": The monitoring of newsgroup and forums, blogs and social media is easily possible by sentiment analysis. Opinion mining and sentiment analysis can automatically detect arrogant words, over heated words or hatred language used in emails or forum entries or tweets on various internet sources.
- Opinion spam detection: Since internet is available to all, anyone can put anything on internet, this increased the possibility of spam content on the web. People may write spam content to mislead the people.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Opinion mining and sentiment analysis can classify the internet content into' spam' content and 'not spam' content.

- Policy Making: Through Sentiment analysis, policy makers can take citizen's point of view towards some policy and they can utilize this information in creating new citizen friendly policy.
- Decision Making: People's opinion and experience are very useful element in decision making process. Opinion mining and Sentiment analysis gives analyzed people's opinion that can be effectively used for decision making.

VII. RESEARCH CHALLENGES IN OPINION MINING AND SENTIMENT ANALYSIS

The main challenges that are faced by Opinion mining and sentiment analysis are the following:

7.1 Detection of spam and fake reviews

The web contains both authentic and spam contents. For effective Sentiment classification, this spam content should be eliminated before processing. This can be done by identifying duplicates, by detecting outliers and by considering reputation of reviewer.

7.2 Limitation of classification filtering

There is a limitation in classification filtering while determining most popular thought or concept. For better sentiment classification result this limitation should be reduced. The risk of filter bubble gives irrelevant opinion sets and it results false summarization of sentiment.

7.3 Asymmetry in availability of opinion mining software

The opinion mining software is very expensive and currently affordable only to big organizations and government. It is beyond the common citizen's expectation. This should be available to all people, so that everyone gets benefit from it.

7.4 Incorporation of opinion with implicit and behaviour data

For successful analysis of sentiment, the opinion words should integrate with implicit data. The implicit data determine the actual behaviour of sentiment words.

7.5 Domain-independence

The biggest challenge faced by opinion mining and sentiment analysis is the domain dependent nature of sentiment words. One features set may give very good performance in one domain, at the same time it perform very poor in some other domain.

7.6 Natural language processing overheads

The natural language overhead like ambiguity, co-reference, Implicitness, inference etc. created hindrance in sentiment analysis too.

VII. RESEARCH SCOPE IN OPINION MINING AND SENTIMENT ANALYSIS

The major research scope areas in sentiment analysis are:

- Spam Detection Sentiment Analysis;
- Sentiment Analysis on short Sentence like abbreviations;
- Improving sentiment word identification algorithm;
- Developing fully automatic analyzing tool;

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

- Effective Analysis of policy opinionated content;
- Successful handling of bi polar sentiments;
- Generation of highly content lexicon database.

VIII. CONCLUSION

Thus, Opinion Mining and Sentiment analysis has wide area of applications and it also facing many research challenges. Since the fast growth of internet and internet related applications, the Opinion Mining and Sentiment Analysis become a most interesting research area among natural language processing community. A more innovative and effective techniques required to be invented which should overcome the current challenges faced by Opinion Mining and Sentiment Analysis.

REFERENCES

[1] Liu, B. (2010), "Sentiment Analysis and Subjectivity". Appeared in Handbook of Natural LanguageProcessing, Indurkhya, N. &Damerau, F.J. [Eds.].

[2] Dave K., Lawrence, S. &Pennock, D.M. (2003), "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews". In Proceedings of the 12th International Conference on World Wide Web, p. 519- 528.

[3] Hu, M. &Lui, B. (2004), "Mining and Summarizing Customer Reviews". In Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2004 (KDD-2004), p. 168–177.

[4] Pang, B. & Lee, L. (2004), "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts". In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, p. 271-278.

[5] Pang, B., Lee, L. &Vaithyanathan, S. (2002), "Thumbs Up? Sentiment Classification Using Machine Learning Techniques". In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2002), p. 79-86.

[6] Boiy, E., Hens, P., Deschacht, K. &Moens, M.-F. (2007), "Automatic Sentiment Analysis in OnLine Text". In Proceedings of the Conference on Electronic Publishing (ELPUB-2007), p. 349-360.

[7] Segaran, T. (2007), Programming Collective Intelligence. Sebastopol: O'Reilly Media, Inc.

[8] Pang, B. & Lee, L. (2008), "Opinion Mining and Sentiment Analysis". In Foundations and Trends in Information Retrieval 2 (1-2), p. 1–135.

[9] Gamon, M. (2004), "Sentiment Classification on Customer Feedback Data: Noisy Data, Large Feature Vectors, and the Role of Linguistic Analysis". In Proceedings of the International Conference on Computational Linguistics (COLING 2004), p. 841-847.