An Overview Of Opinion Mining

Mrs. A Hema[1] MCA, Mphil,PGDBI, Associate Professor & Head , Department of Computer Applications (UG) , Kongunadu Arts and Science College, Coimbatore-29

M. Abhayadev[2], Mphil Scholar, Kongunadu Arts and Science College, Coimbatore-29

Abstract

Opinion Mining is the field of study that analyzes people's opinions, sentiments, evaluations, attitudes and emotions from written language **[9]**. Opinion mining is also called as Sentiment Analysis; it is one of the most active research areas in natural language processing and is also widely studied in Data Mining, Web Mining, and Text Mining. This research has spread outside of Computer Science to the Management Sciences and Social Sciences due to its importance to business and society as a whole. The growing importance of Sentiment Analysis coincides with the growth of social media such as reviews, forum discussions, blogs, micro-blogs, twitter, and social networks. For the first time in human history, we now have a huge volume of opinionated data recorded in digital form for analysis. Sentiment Analysis systems are being applied in almost every business and social domain because opinions are central to almost all human activities and are key influencers of our behaviors. Our beliefs and perceptions of reality, and the choices we make, are largely conditioned on how others see and evaluate the world. For this reason, when we need to make a decision we often seek out the opinions of others. This opinion is helpful not only for individuals but also for organizations.

I. Introduction

Data mining is a process of extracting information from big data set and transforms it in to an understandable structure for further use **[11]**. Data mining identifies a problem, along with collection of data that can lead to better understanding and Computer models to provide statistical or other means of analysis. This may be supported by visualization tools, that display data, or through fundamental statistical analysis, such as correlation analysis. Data mining is expanding rapidly, with many benefits to business. The most profitable application areas have been to identify the greater probabilities of responding to marketing media. Banks are more accurately predict the likelihood of people to respond different services offered. Many companies are using this technology to identify their blue-chip customers so that they can provide them the service needed to retain them. Data mining have the number of applications like Market segmentation and Customer churn. The Market segmentation is behaviors that are common among the customers. The Customer churn will allow to estimate which customers are the most likely to stop purchasing the products or services and go to other competitors **[10]**. In addition to this, a company can use data mining to find out which purchases are the most likely to be fraudulent.

Data mining can allow becoming better at what we need. It is also a powerful tool for those who deal with finances. A financial institution such as a bank can predict the number of defaults that will occur among their customers within a given period of time, and they can also predict the amount of fraud that will occur as well. Another potential application of data mining is the automatic recognition of patterns that were not previously known. A tool which could automatically search the database to look for patterns which are hidden. Assume there is a provision to access this technology. It would be able to find relationships which allow making strategic decisions **[8]**.

Opinion mining is the key to find actionable information, or information that can be utilized in a concrete way to improve profitability **[10].** Some of the earliest applications were in retailing, especially in the form of market basket analysis. Generally speaking, Sentiment Analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. The attitude may be his or her judgment, affective state (that is to say, the emotional state of the author when writing), or the intended emotional communication (that is to say, the emotional effect the author wishes to have on the reader). Knowledge Discovery and Data Mining (KDD) is an interdisciplinary area focusing upon methodologies for extracting useful knowledge from data. The ongoing rapid growth of online data due to the Internet and the wide spread use of databases have been created an immense need for KDD methodologies. The challenge of extracting knowledge from data draws upon research in statistics, databases, pattern recognition, machine learning, data visualization, optimization, and high-performance computing, to deliver advanced business intelligence and web discovery solutions **[9].**

II. Literature Study

Opinions are different from facts. A fact can be defined as something said to have happened or supposed to true. However as a mining person, it needs to know how reliable statements are before they can report them as facts. This determines how we present it in the final results. The facts are the objective expressions which describe the entities, events and properties. An opinion is a conclusion reached by someone after looking at the facts. Opinions are based on what people believe to be facts **[10].** This can include probable facts and even probable lies, although few people will knowingly give an opinion based on a proven lie the opinion is the subjective expression which describes people's opinions, emotions and sentiments towards entities and their properties. The opinion is important that whenever we need to make a decision to hear other's opinions, so all the individuals can express their opinion through reviews, forums, discussions, and blogs **[11]**. There are two types Opinions, Explicit opinion and Implicit Opinion in explicit opinion an attribute explicitly expressed in a subjective sentence, for example the voice quality of this phone is amazing. In implicit opinion an attribute implied in an objective sentence for example the headset broke in two days.

The web has dramatically changed the way that people express their opinion; they can now post reviews of products at merchant sites and express their views on almost anything in internet forums, discussions groups, blogs etc. business always want to find public or consumer opinion on their product and services. Potential customers also want to know the opinions of existing users before they use a service or purchase a product. For the recent years there will be a tremendous increase in demand for Sentiment Analysis tools by companies willing to monitor people's opinions of the company

and on its products and services but also by social science researchers. To fulfill the increasing demands for such tools, more and more researchers and companies are releasing products to perform sentiment analysis, many of them claiming to be able to perform sentiment analysis of any type of document in every domain. Unfortunately, some experience shows that Sentiment Analysis tools also working across domains which does not yet exist. The main reason Sentiment Analysis is so difficult is that words often take different meanings and are associated with distinct emotions depending on the domain in which they are being used. The use of a word like fingerprints may represent a major breakthrough in a criminal investigation but a major headache for smart phone manufacturers **[10].**Opinion Mining draws on computational linguistic, information retrieval, text mining, natural language processing, machine learning, statistics and predictive analysis **[11].**

A basic task in Sentiments Analysis is classifying the polarity of a given text at the sentence, whether the expressed sentence is positive, negative, neutral or Sentiment classification looks, for instance, at emotional states such as angry sad and happy. There are so many applications based on Opinion Mining in this only three are very important Business, Individual and Ads placement applications, In Business application spending huge amount of money to find Consumer Sentiments and Opinions. In individual side application a person interested in other's opinion when purchasing a product lastly Ads placement application in this placing an ads in the user generated content. There are three components in Opinion Mining are Opinion holder, Object and Opinion. The Opinion holder means who is talking, Object is Item on which opinion is expressed and Opinion stands Attitude or view of opinion holder

There are several challenges in Opinion Mining the One, is a word that is considered to be positive in one situation may be considered negative in another situation. If a customer said a laptop's battery life was long, that would be a positive opinion. If the customer said that the laptop's start-up time was long, however, that would be is a negative opinion. These differences mean that an opinion system trained to gather opinions on one type of product or product feature may not perform very well on another. A second challenge is that people don't always express opinions the same way. Most traditional text processing relies on the fact that small differences between two pieces of text don't change the meaning very much. In Opinion Mining, however, the movie was great is very different from the movie was not great finally people can be contradictory in their statements. Most reviews will have both positive and negative comments, which is somewhat manageable by analyzing sentences one at a time.

III. Conclusion

Data mining is a process of extracting information from big data set and transforms it in to an understandable structure for further use. Opinion Mining is derived using different data mining rules and technique so Data Mining and Opinion Mining are very closely inter connected. In business side they consider public or consumer's opinions for their product and services. Data mining is the Computer-assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data. Data mining tools predict behaviors and future trends, allowing businesses to make proactive knowledge and driven decisions.

An important part of our information-gathering behavior has always been to find out what other people think. With the growing availability and popularity of opinion-rich resources such as online review sites and

personal blogs, new opportunities and challenges arise as people now can, and do, actively use information technologies to seek out and understand the opinions of others. The sudden eruption of activity in the area of Opinion Mining and Sentiment Analysis, which deals with the computational treatment of opinion, Sentiment, and subjectivity in text, has thus occurred at least in part as a direct response to the surge of interest in new systems that deal directly with opinion mining as so importance in our world**.**

Mining is of different forms in which Text mining usually involves the process of structuring the input text (usually parsing, along with the addition of some derived linguistic features and the removal of others, and subsequent insertion into a database), deriving patterns within the structured data, and finally evaluation and interpretation of the output. Image mining deals with the extraction of image patterns from a large collection of images. The Mining is moving to World Wide Web towards a more usefully environment in which users can quickly and easily find the information they need. Using search engine user can find the knowledge through keyword and hyperlinks, so to fulfill the increasing demands for such tools, more and more researchers and companies are releasing products to perform Sentiment analysis, many of them claiming to be able to perform Sentiment analysis of any type of document in every domain. Unfortunately, some experience shows that Sentiment analysis tools also working across domain which does not yet exit. However, Opinion Mining also called as sentiment analysis widely spread outside of Computer Science with the growth of Social media such as reviews, forms, Social network due to its important both Society and Business.

References

1. Lippmann, R. P., Pattern classification using neural networks, IEEE Communications Magazine, pp. 47–54. (1989)
2. Lu, H., Setiono, R.Liu.H., Effective data mining using neural networks, IEEE Transactions on Knowledge and Data Engineering,48-56, (1991).
3. R. Agrawal, A. Swami., Database mining: A performance perspective, IEEE Transactions on Knowledge and Engineering 5:6, 914–925, (1993).
4. Agrawal R, Imielinski T and Swami A. May, Mining association rules between sets of items in large databases. Proc. 1993 Int Conf. Management of Data (SIGMOD-93), 207-216, (1993).
5. R. Agrawal, R. Srikant., Fast algorithms for mining association rules, Proceedings of the 20th International Conference on Very Large Data Bases *(VLDB'94)* Santiago, Chile, (487–499), (1994).
6. Kennedy, R. L., Lee, Y.Roy, B. V, Reed, C. D., and Lippmann, R. P., Solving data mining problems through pattern recognition, Prentice Hall, (1998).
7. H.K. Bhargava., Data mining by decomposition).Adaptive search for hypothesis generation, *INFORMS* Journal on Computing, 11:3, (239–247), (1999).
8. Perer D Turney., Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews, ACL, (2002).
9. Bo Pang , Lillian Lee , Shivakumar vaithyanathan., Thumbs up Sentiment Classification using Machine Learning *Techniques*, CoRR cs.CL/0205070, (2002).
10. Charlotta Engstrom., Topic Dependence in Sentiment Classification, Master's thesis, University of. Cambridge, (2004 July).
11. Dr. N. Karthikeyani Visalakshi.,Associate Professor,Department of Computer Applications,Kongu Engineering College,Perundurai .Research Challenges in Opinion Mining NATIONAL SEMINOR ON RESENT TRENDS IN DATA MINING held in Vaishiya College of arts and science, (2013).