

AN EXPLORATION OF TEXT MINING TECHNIQUES WITH EXPLOIT OF SENTIMENT ANALYSIS

¹Dr. B.Jagadhesan, MCA. MBA. M.Phil., PhD, ²R.Georgina Shefani.
¹Associate Professor & Head of the Department of Computer Science,p
 Dhanraj Baid Jain College,
 Chennai, India,
 Email:bjagadhesan@gmail.com

²M.Phil Scholar Department of Computer Science Dhanraj Baid Jain College,
 Chennai, India,
 Email:georgina.shefani@gmail.com

ABSTARCT

Sentiment analysis is a text analysis method that detects polarity (e.g. a positive or negative opinion) within text, whether a whole document, paragraph, sentence, or clause. Understanding people's emotions is essential for businesses since customers are able to express their thoughts and feelings more openly than ever before. By automatically analyzing customer feedback, from survey responses to social media conversations, brands are able to listen attentively to their customers, and tailor products and services to meet their needs.

Keywords-Data Mining, Sentiment Analysis, Text Mining Algorithm, KnowledgeDiscovery, Applicaions.

I. INTRODUCTION

Sentiment analysis is the process of using natural language processing, text analysis, and statistics to analyze customer sentiment. Any business has to analyze and understand the sentiment of their customers on what people are saying, how they're saying it, and what they mean to be the best in their own field. The sentiment of the customers using a particular brand can be found in tweets, comments, reviews. In sentiment analysis we understand the emotions of these people through the use of software, and it is a much necessity for both developers and business leaders pertaining to a modern workplace. As with many other fields, advances in deep learning have brought sentiment analysis into the foreground of cutting-edge algorithms. Today we use natural language processing, statistics, and text analysis to extract, and identify the sentiment of words into positive, negative, or neutral categories. It's estimated that 80% of the world's data is unstructured, in other words it's unorganized. Huge volumes of text data (emails, support tickets, chats, social media conversations, surveys, articles, documents, etc), is created every day but it's hard to analyze, understand, and sort through, not to mention time-consuming and expensive. Sentiment analysis, however, helps businesses make sense of all this unstructured text by automatically tagging it.

II. TYPES OF SENTIMENT ANALYSIS

1. Standard Sentiment Analysis

This is the most popular used type in the standard sentiment analysis as it recognizes and studies the expression and the tone of opinions about a product or service. It helps to categorize the subjective details whether they are positive, negative or neutral data. For instance:

- 'I love Product A particularly for its performance' – Positive
- 'I still need time to operate Product A according to its various functions' – Neutral
- 'Too many buttons and functions on Product A, so confusing!' – Negative

2. Fine-Grained Sentiment Analysis

This is another type that focuses on recognizing people's opinions but it provides a more precise output. The feedback gathered is ranked to 5-star ratings, which are:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

The data is usually mapped to 5 stars as 'Very Positive' to 1 star as 'Very Negative'. Besides the mentioned common terms, other traits like feelings can also be used to measure whether the responses are positive or negative. Words like 'anger', 'sadness' and 'worries' are associated with negative sentiment, while words like 'happiness', 'love' and 'enthusiasm' are linked with the positive sentiments.

3. *Emotion Detection*

This type of sentiment detection finds the feelings of the customer with the response that is been provided. The words in the responses are associated with various different feelings such as happiness, frustration or anger and many more. The only downside of this type is how the customers express their opinion and the choice of words. For example, a customer might say ‘the product is killing me’ or he might say ‘the graphics is killing it’, whereas the word ‘kill’ is used in both the sentence and it can be either positive or negative.

4. *Aspect-Based Sentiment Analysis*

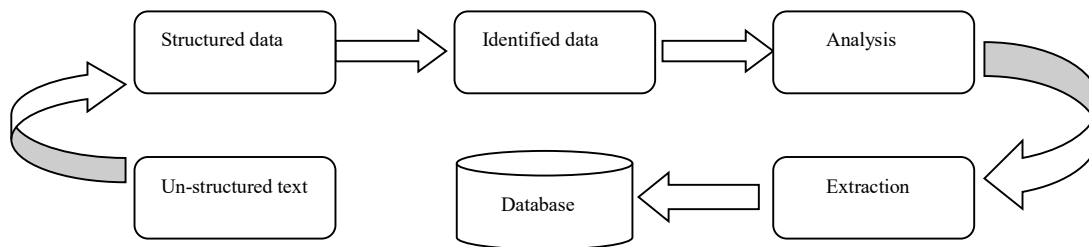
Product reviews are considered a very good way to understand the customers insight about the product. But by discussing the certain aspects of the products or specific details it might be useful for the companies. For example, ‘That laptop has a great graphics card but a very low sound effect ‘

5. *Intent Analysis*

Using sentiment analysis, it is very easy to solve customer problems or to offer customer support. Here Intent analysis tries to understand the customer issues from the reviews that is given a try to solve them to get a better customer satisfaction. For example: ‘I can’t take it anymore! The XX application kept closing whenever I log in. Can you help?’Using sentiment analysis, it is very easy to detect customers’ complaints and resolve problems.

III. TECHNIQUES USED IN TEXT MINING

Text mining is an essential field that employs collective methods and approaches from many areas, such as



1. *Information Retrieval*

Information retrieval is an old-style area of reaserch.IR refers to the process of extracting relevant and associated patterns based on a specific set of words or phrases that is been inputted. In this text mining technique, IR systems make use of different sets of algorithms to track and monitor user behaviours and discover relevant data accordingly. Google and Yahoo search engines are the two most renowned IR systems. The rise of the World Wide Web has enhanced the thought and the need for more refined search.

2. *Extraction of Information*

This is the most famous text mining technique. The process of extracting meaningful information from vast loads of textual data is called extraction of information. Identifying the extraction of entities, attributes, and their relationships from semi-structured or unstructured texts is the main focus of text mining. All the information is been extracted is then stored in a database for future access and retrieval. The efficiency and relevancy of the outcomes are checked and then carefully evaluated using precision and recall processes.

3. *Summarization of Text*

Text summarisation refers to the process of automatically creating a compressed version of a specific text that holds some valuable information for the end-user. The main aim of text mining technique is to browse through multiples of text sources to craft summaries of texts containing a considerable proportion of information in a minimalistic format, keeping the overall meaning and intent of the original documents essentially the same. Text summarisation integrates and combines the various methods that employ text categorization like decision trees, neural networks, regression models, and swarm intelligence.

4. *Categorization of Text*

The essential feature of text mining is text categorization. It is a supervised process and uses a predefined document based on their content. The categorization provides to find exactly which domain category in use, a defined text file is retransmitted. Using categorization tools, systems can assign unpublished documents to the most accurate category available, based on the classification or specific topic. It is a set of text documents, the method to find the topic or the precise topics for each document. Nowadays, the automatic categorization of texts is applied in a variety of contexts, from automatic or semi-automatic indexing of texts to the delivery of personalized spots, spam filtering and categorization of Web pages in hierarchical catalogs, automatic generation of metadata and text type detection, monitoring topic and many others.

5. Clustering of Text

Clustering is one of the most crucial text mining techniques. It seeks to identify intrinsic structures in textual information and organize them into relevant subgroups or clusters for further analysis. A significant challenge is to form meaningful clusters from the unlabelled textual data without having any prior information on them. Cluster analysis is a standard text mining tool that assists in data distribution or acts as a pre-processing step for other text mining algorithms running on detected clusters.

Techniques	Characteristics
Retrieval of text	Retrieval of valid and valuable information from the unstructured data
Extraction of data	Extracting information from structures data
Summarization	Trimming the length of a sentence by retaining the essence and meaning of the same
Categorisation	Categorisation based on the documents
Clustering	Collection of documents in clusters e with further classification and analysis of the text documents

IV. BENEFITS OF SENTIMENT ANALYSIS

1. Improve Customer Service

One of the benefits of sentiment analysis is being able to track the key messages from customers' sentiments and opinions about a brand. This helps the customer service team to be aware of all the issues or problems faced by them. As this method allows the organizations to understand their customers better, sentiment analysis provides a clear picture of the problems faced by them and persuades the organisation to look for a better solution. Plus, having quick sentiment analysis detection over customer unfavourable remarks, the organisation can act swiftly by investigating the root cause and provide the customer service department an effective resolution. Nothing beats a prompt response in addressing an issue from the company itself.

2. Develop Quality Products

Customer satisfaction is crucial and a very important deciding factor, customers remain loyal only when the products gives them 100% satisfaction. Hence another of the benefits of sentiment analysis make the whole process easier and at the same time providing opportunities that tends to the improvement of the same. This allows the marketing team to research about the current trends and customers' preferences better by directly or indirectly conversing with them. The responses from the customers can be used as the guideline to improve the service quality, better future product development, reduce customer churn or improve how the product is presented.

3. Discovering New Marketing Strategies

The more data and information gathered through sentiment analysis, the organizations could develop an effective and appealing marketing strategy. The outcome from the strategies that been followed can be measured from the customers' positive or negative messages received by customer. By observing the customers' conversations and thread on their social media and detect the specific messages related to that brand, specific marketing campaigns can be designed to target consumers.

4. Improve Media Perceptions

Another important factor of Sentiment Analysis is to be able to track the mentality of the journalists, writers, columnists, market analysts, media researchers or independent contributors towards the particular company, be it a product, service, company values, human resources etc. This is crucial because any misinterpretation or negative approach can lead to negative key messages which forms an undesirable perception for any company. Knowing what is being written by who is very important because certain subjects are appealing content for the people but also creates a negative image for the particular brand

5. Increasing Sales Revenue

Sentiment analysis captures the intentions and moods of the customers of a brand and this is definitely a great way to improve sales profits. As negative review or thread that is found, the marketing team works their magic to solve the problems and optimize the product quality, organizations can estimate a higher monetary return. By making use of sentiment analysis the

management improves the products and services. In addition, when there is a huge customer satisfaction the customers feel that they are being heard and their needs are taken care of, thus improving a company's image as well.

6. Improve Crisis Management

Monitoring the responses of the customers or opinions towards any brand would definitely help in identifying any issues in a quickly manner. Avoiding any pending complaints is one of the main purposes of sentiment analysis which allows efficient crisis management. Timely preventive steps are very important because these prevent communication crisis. Companies have a sharp eye on the internet to watch out for negative threads or comments and so any issues can be dealt before the issue is escalated.

V. CONCLUSION

Text mining deals with texts in natural language stored in semi-structured or unstructured formats. The document dealt with various text mining techniques, their applications in various fields and the comparison of different text mining techniques, which can be further improved. The availability of large volume of text-based data needed to be analyzed to obtain useful information. Different type of sentiment analysis can be used to various needs to get the desired output and prediction. Selection and use of correct methods and tools according to the domain will help to make the text mining method easier and more efficient. Domain knowledge-based integration, concepts-based granularity, multilingual text of refinement, and using natural language processing ambiguity are major issues and risks that arise during text mining techniques. By using the appropriate techniques for the particular set of data will give us more benefits. More than any other field, text mining is of great advantage in life science and health care

REFERENCES

- [1] S. M. Weiss, N. Indurkha, T. Zhang, and F. Damerau, Text mining: predictive methods for analyzing unstructured information. Springer Science and Business Media, 2010.
- [2] S.-H. Liao, P.-H. Chu, and P.-Y. Hsiao, "Data mining techniques and applications—a decade review from 2000 to 2011," Expert Systems with Applications, vol. 39, no. 12, pp. 11 303–11 311, 2012..
- [3] K. Sumathy and M. Chidambaram, "Text mining: Concepts, applications, tools and issues-an overview," International Journal of Computer Applications, vol. 80, no. 4, 2013.
- [4] N. Padhy, D. Mishra, R. Panigrahi et al., "The survey of data mining applications and feature scope," arXiv preprint arXiv:1211.5723, 2012.
- [5] Juan Jose Garcia Adeva and Rafael Calvo, "Mining Text with Pimiento", University of Sydney.
- [6] Rashmi Agrawal, Mridula Batra, "A Detailed Study on Text Mining Techniques", IJSCE, ISSN: 2231-2307, Vol. 2, Issue-6, January 2013.
- [7] David Osimo and Francesco Mureddu, "Research Challenge on Opinion Mining and Sentiment Analysis".
- [8] Dongjoo Lee et al, "Opinion Mining of Customer Feedback Data on the Web". Seoul National University.