



# A Review on Sentimental Analysis on Facebook Comments by using Data Mining Technique

**Rupinder Kaur**

Research Scholar, Punjabi University, Patiala, [dhaliwalruby12@gmail.com](mailto:dhaliwalruby12@gmail.com)

**Dr. Harmandeep Singh**

Assistant Professor, Punjabi University, Patiala, [harmanjhajj@yahoo.co.in](mailto:harmanjhajj@yahoo.co.in)

**Dr. Gaurav Gupta**

Assistant Professor, Punjabi University, Patiala, [gaurav.shakti@gmail.com](mailto:gaurav.shakti@gmail.com)

**ABSTRACT:** *Data mining is the investigation periods of the “data discovery in documents”. This is a method for deciding plans and extracting the information from huge set of data. It is the procedure of mining knowledge from data Sentiment analysis refers to the use of natural language processing [4]. Sentiment analysis, also called opinion mining, is the field of study that analyzes people’s sentiments, opinions, and emotions towards entities [7]. These entities might be a thing or a film, surveys of people, products, issues and topics that truly matters. Social sites for example Facebook and Twitter are that, where characters put their status or sentiments. People comment on their facebook account concerning any correct subject of their consideration [3].*

**Keywords:** *Data mining, sentimental analysis, facebook comments, classification, slang words*

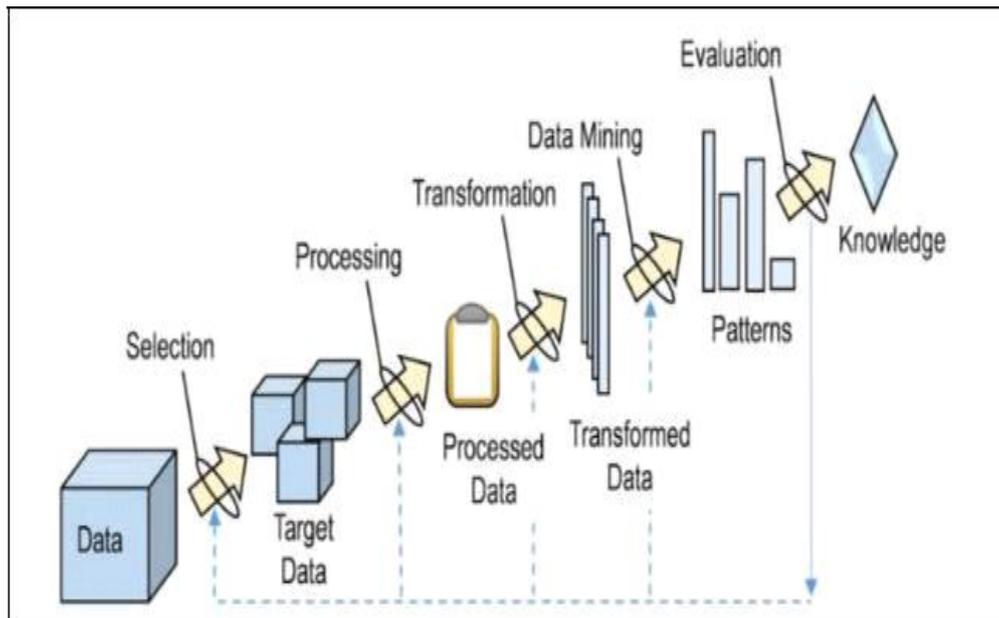
## I. INTRODUCTION

Data mining, also called knowledge discovery in databases, is the process of discovering interesting and useful patterns [17]. The aim of data mining process is to gather information from a data set and convert it into understandable form [18].

### 1.1 Process of data mining or knowledge discovery in database

Knowledge discovery in database is the method of extracting useful information from subordinate databases. Knowledge discovery in database or data mining involves the steps like data processed from data selection, interpretation, cleansing, transforming the raw data into some information, integrating and evaluating the pattern for that information received. Following are the overview of stages in the data identification process [10].

- **Data selection:** Data come from various resources are combined as single entity data called target data.
  - **Data cleaning:** Data from target data are pre-processed and transformed into processed data.
  - **Data transformation:** Modification of the information into the shape that is required for mining operations is called information change.
  - **Data mining:** The result evolved is been transformed to form some patterns by applying some data mining algorithms.
  - **Pattern evaluation:** Finally those results are incorporated to form some useful information called knowledge.
- Figure 1 shows the process of data mining or knowledge discovery in database



### 1.2 Data mining techniques: Following are often used techniques in data mining.

- **Classification:** It is the most commonly applied data mining technique. But it requires sat of pre-defined examples to create a model and applied on large datasets [10].
- **Support vector machine:** These are supervised learning models with associated learning algorithms that analyses data used for regression analysis and classification [10].
- **Decision tree:** Decision tree is a graph of decisions and their possible significances, represented in the form of branches and nodes. A decision tree contains a root node, branches, and leaf nodes [10].
- **Regressions:** A task used to predict the number such as age, height, income distance etc [10].
- **Prediction:** It is a method used in predicting the outcome based on available data by incorporating with unavailable data sets of future [10].

### 1.3 Introduction to sentiment analysis

Social media has given web users a venue for expressing and sharing their thoughts on different events [5]. Sentiment analysis is a method of computing and satisfying a view of a person given in a piece of a text, to identify persons thinking about any topic is positive negative or neutral [2].

Sentiment analysis refers to the use of natural language processing [4]. Sentiment analysis, also called opinion mining, is the field of study that analyzes people’s sentiments, opinions, and emotions towards entities [7]. These entities might be a thing or a film, surveys of people, products, issues and topics that truly matters.

Social sites for example Facebook and Twitter are that, where characters put their status or sentiments. People comment on their facebook account concerning any correct subject of their consideration [3].

**1.4 Application areas:** Following are the different regions where conclusion examination can be utilized:

- **E-trade:** Many sites give outline of their items that enable clients to present these perspectives are useful for both different clients and the item producers [9].
- **Marketing:** It encourages the item build to discover which clients are faithful and which are not and how to make new clients their dedicated clients [11, 12]
- **Politics:** Before the real decision result the perspective of open can be dividing by their remarks or surveys on the online networking. A number of voting applications are accessible in market to break down the perspective of open [12].
- **Decision making:** Opinion mining and sentiment analysis gives analyzed people's opinion that are used for decision making [11].

## II. LITERATURE REVIEW

During the age of time, reading certain of the research papers has been done which is summarized as below:-

**Kumar *et al.* (2019).[1] “ Sentiment Analysis of Electronic Product Tweets Using Big Data Framework”** discussed about the different sale tweets used to examine the sentiments of customers regarding electronic goods. The experimental results of the proposed work will be useful for various business companies to take business verdicts, which will further improve the product sales. In the current scenario, millions of tweets are produced by people every year. But handling these huge unstructured tweets is not possible through the traditional platform. Therefore, big data framework, such as Hadoop and Spark, is used to grip such kind of large data.

**Mathapati *et al.* (2016).[2] “Sentiment Analysis and Opinion Mining from Social Media : A Review”** discussed about the need for automated analysis techniques to extract sentiments and opinions sent in the user-comments. words provide fine-grained analysis on the customer reviews. The labeled data required for training a classifier is expensive and hence to overcome, Domain Adaptation technique is used. In this technique, Single classifier is designed to classify homogeneous and heterogeneous input from di\_erent domain. This paper focuses on the survey of the existing methods of Sentiment analysis and Opinion mining techniques from social media.

**Rastogi *et al.* (2014).[3] “A Sentiment Analysis based Approach to Facebook User Recommendation”** discussed about system to offer new friends who have similar interests but having different opinions. The motivation of this work is that users may share similar interests but have dissimilar opinions on them. In this paper, a user recommendation technique based on a novel weighting function is proposed, which consider not only user interests, but also his sentiments. Such function allows to build better commendation process than other content-based approaches. First round results based on a comparative analysis show the benefits of the novel approach in comparison with some modern user recommender systems.

**Gupta *et al.* (2017).[4] “Sentiment Analysis of Twitter and Facebook Data Using Map-Reduce”** discussed about Twitter and Facebook's amusing source of data for opinion mining or sentiment analysis and this vast data can be used to find the sentiments of people on a specified topic or product. In this paper, system is proposed which involves collecting data from social network using the Twitter and Facebook API's. Then, the challenges of big data are answered using Hadoop through map reduce framework where the complete data is mapped and reduced to smaller sizable data to ease of handling and finally contains analyzing the collected data and represent the results through graphs.

**Isah *et al.* [5] “Social Media Analysis for Product Safety using Text Mining and Sentiment Analysis”** discussed about user created content from social media platforms that can provide early clues about product allergies, adverse events and product counterfeiting. This paper accounts a work in progress by means of contributions including: the growth of a framework for assembling and analyzing the views and experiences of users of drug and cosmetic goods using machine learning, text mining and sentiment analysis; the application

of the projected framework on Facebook comments and data from Twitter for brand analysis, and the description of how to develop a product safety lexicon and training data for modeling a machine learning classifier for drug and cosmetic product sentiment prediction.

**Gürsoy *et al.* (2017)[6]. “Social Media Mining and Sentiment Analysis for Brand Management”** discussed that Corporate companies want to gain from big data studies extra. Although it affects different company dynamics in various areas, especially social media services have become very significant for the marketing and CRM departments of businesses. In this way, communication is always recognized with the customers and the use of Big data in these fields is seen as one of the utmost important steps of the firms in becoming a big trademark. In this study, social mass media and digital data of the 3 major firms operating in the manufacture, technology and food industry in Turkey were examined. The data was attained with the help of API and Web Crawler.

**Nigam *et al.* (2015)[7]. “Opinion Mining of Online Shopping Sites through Facebook pages using Graph API and FQL Query”** discussed about posting our particulars and adding and refusing people as friends and creating so called a social virtual human network which shares ideas, feeling, status, pictures and so on. This paper presents the hands-on implementation of how to know the thorough status of social networking Facebook account so as to see the total number of people on our net, our friends, their records gender-wise, their status, their IDs, picture shared, likes unlike, different responses of people in our network on our posts, all can be calculated as a summary with the aid of Analysis Software R with an additional package of Facebook R on it.

**Zamani *et al.* [8]. “Sentiment Analysis: Determining People’s Emotions in Facebook”** discussed to identify the opinion mining and sentiment analysis apparatuses for extracting both English and Malay words in Facebook. This effort begins with converting unstructured information into meaningful lexicons after extracting the Facebook's contents and stored in a database after physical identifications are carried out. With sentiment analysis, emotions are categorized into happy (positive), unhappy (negative) and emotionless. As a case study, a problem on an inspection results is posted and results of students' responses are resolute. This study is momentous of enabling the stakeholders such as administrators and businessmen to monitor any discussion dispute for enhancing their facilities.

**Patil and Thakare (2017)[9]. “Analyzing Public Sentiment Variations on Twitter and Facebook”** discussed about interchange of views, ideas, expressions, feelings and opinions on social networking sites like Twitter and Facebook. In this work, analyzation of public sentiment variations in an explicit time period about a explicit target on Twitter and Facebook both is done. This kind of analysis is helpful in various fields for taking proper conclusions and deteriorating public opinion.

**Salloum *et al.* (2017)[10]. “A Survey of Text Mining in Social Media: Facebook and Twitter Perspectives”** discussed about a mutual practice to not write a sentence with correct grammar and spelling at social networking sites which leads to diverse kinds of uncertainties like lexical, syntactic, and semantic and due to this type of uncertain data, it is inflexible to find out the actual data order. This study aims to describe how studies in social media have used text analytics and text mining methods for the purpose of categorizing the key themes in the information. This survey concentrated on examining the text mining studies related to Facebook and Twitter; the two overriding social media in the world.

### III. Conclusion

Sentiment analysis can be seen as a utilization of content order. The primary occupation of content gathering is how to stamp writings with a predefined set of gatherings. Content gathering has been helpful in different zones for example, article ordering and content cleaning. Huge quantities of comments or surveys are posted by people in general every day. So to distinguish the assessment of open towards a particular post is by physically analyze and discover each comment.

People use very awkward words to express their feelings & most of the people use shortcuts e.g. osm for awesome, lol for laughing out loud & many more, so this is sometime creating difficulty for the person who is not familiar with these words and cannot recognize the sentiments of the person.

## REFERENCES

- [1]. Gupta, J. Pruthi, N. Sahu (2017), “**Sentiment analysis of tweets using machine learning approach**”, International Journal of Computer Science and Mobile Computing, **6(4)**, pp. 444-458.
- [2]. P. Jain, M. V. D. Katkar (2015), “Sentiments analysis of Twitter data using data mining”, International Conference on Information Processing, **10**, pp.807-810.
- [3]. P. Rajan, S. P. Victor (2014), “**Web sentiment analysis for scoring positive or negative words using tweeter data**”, International Journal of Computer Applications, **96(6)**, pp. 33-37.
- [4]. Alvares, N. Thakur, S. Patil, D. Fernandes, K. Jain (2016), “**Sentiment analysis using opinion mining**”, International Journal of Engineering Research & Technology, **5(4)**, pp.88-91.
- [5]. M. Bandgar, D. S. Sheeja (2016), “**Analysis of real time social tweets for opinion mining**”, International Journal of Applied Engineering Research, **11(2)**, pp. 1404-1407.
- [6]. S. Dattu, P. Deipali, V.Gore (2015), “**A survey on sentiment analysis on Twitter data using different techniques**”, International Journal of Computer Science and Technologies, **6(6)**, pp.5358-5362.
- [7]. Chandni, N. Chndra, S. Gupta, R. Pahade (2015), “**Sentiment analysis and its challenges**”, International Journal of Engineering Research & Technology, **4(3)**, pp.968-970.
- [8]. E. Olearly (2015), “**Twitter mining for discovery, prediction and causality: applications and methodologies**”, International Journal of Intelligent Systems in Accounting and Finance Management, **22(3)**, pp.222-247.
- [9]. G. Hu, P. Bhargava, S. Fuhrmann, S. Ellinger, N. Spasojevic (2017), “**Analyzing users’ sentiment towards popular consumer industries and brands on twitter**”, International Conference on Data Mining Workshops, 381-388.
- [10]. G. Sabarmathi, D. R. Chinnaiyan (2017), “Reliable data mining tasks and techniques for industrial applications”, IAETSD Journal for Advanced Research in Applied Sciences, **4(7)**, pp. 138-142.
- [11]. H. P. Rahmath (2014), “**Opinion mining and sentiment analysis- challenges and applications**”, International Journal of Application or Innovation in Engineering & Management, **3(5)**, pp.401-403.
- [12]. Smeureanu, C. Bucur (2012), “**Applying supervised opinion mining techniques on online user reviews**”, Informatica Economică, **16(2)**, pp. 81-91.
- [13]. K. I. Umar, F. Chiroma (2016), “**Data mining for social media analysis: using twitter to predict the 2016 Us presidential election**”, International Journal of Scientific & Engineering Research, **7(10)**, pp.1972-1980.
- [14]. K. Sutar, S. Kasab, S. Kindare, P. Dhule (2016), “**Sentiment analysis: opinion mining of positive, negative or neutral twitter data using hadoop**”, International Journal of Computer Science and Network, **5(1)**, pp. 177-180.
- [15]. L. J. Sheela (2016), “**A review of sentiment analysis in twitter data using hadoop**”, International Journal of Database Theory and Application, **9(1)**, pp.77-86.
- [16]. N. Azam, J., M. Abulaish, S., N. A. H. Haldar (2015), “**Twitter data mining for events classification and analysis**”, International Conference on Soft Computing and Machine Intelligence, **10**.
- [17]. P. Tripathi, S. K. Vishwakarma, A. Lala (2015), “**Sentiment analysis of english tweets using rapid miner**”, International Conference on Computational Intelligence and Communication Networks, pp. 668-672.
- [18]. R. Mehta, D. S. Jain (2016), “**Sentiment mining and related classifiers: a review**”, IOSR Journal of Computer Engineering, **18(1)**, pp.50-54.
- [19]. S. A. A. Hridoy, M. T. Ekram, M. S. Islam, F. Ahemed, R. M. Rahman (2015), “**Localized twitter opinion mining using sentiment analysis**”, Decision Analytics, **2(1)**, pp.1-19.
- [20]. Y. Zhao (2013), “**Analysing Twitter data with text mining and social network analysis**”, Proceedings of the 11-th Australasian Data Mining Conference, pp.41-47.