

SENTIMENT ANALYSIS FOR MARKETING

Abstract Sentiment analysis is one of data mining types that estimates the direction of personality's sentiment analysis within natural language processing. Analysing the text computational linguistics are used to deduce and analyse mental knowledge of Web, social media and related references. The examined data quantifies the global society's attitudes or feelings via specific goods, people or thoughts and expose the contextual duality of the knowledge. Sentiment analysis used in different approaches such as products and services reviews. Also is used in healthcare. there is a huge volume of information about healthcare obtainable online, such as personal blogs, social media, and on the websites about medical issues rating that are not obtained methodically. Sentiment analysis provides many benefits such as using medical information to achieve the best result to increase healthcare quality. In this paper, sentiment analysis methods and techniques are presented that used in the medical linguisticKeywords Sentiment analysis Data mining Natural Language Processing (NLP) Computational linguistics

1 Introduction

The online experience gives benefits to the people who are interested in such a product from the other people who wrote their sentiment analysis on such a product [1,2]. There are many sources to find this huge amount of data online, such as social media sites, online forums, personal blogs, etc., including a wide scope of topics [3–6]. People discuss their healthcare cases on a lot of medical websites and forums, and they share their illness, indications, and drugs. The experience of medical centres that people visited, also shares the availability, services, pleasure, etc. [7,8]

It is very important to patients when they learn from other patients' experience to make decisions about their medical issue. Such as choosing hospital, clinic, and medication. This information also benefits the hospitals to know the patients' interests and problems and resolve them. Patients share their experience covered in their own sentiment analysis and passions, which is the power of this type of analysis. [9] has taught sentiment analysis as knowing the sentiments of people about a subject and its features. The medical content that is accessible online is free, in addition to its existing in large volume; therefore, analysing this huge amount of data manually is less effective.

Assessment examination for the most part centered on the programmed acknowledgement of suppositions' extremity, as positive or negative. These days, notion investigation is supplanting the online and customary overview techniques normally led by organizations for finding the general conclusion about their items and administrations to improve their promoting methodology and item notice and help to improve user administration. The online accessibility of enormous content makes it imperative to be broken down. The programmed examination of this data includes a deep comparison of common dialects. Notions and feelings assume a significant role in our day by day lives. They aid basic leadership, learning, correspondence, and circumstance mindfulness in human situations. The significance of handling and understanding vernacular content is expanding because of the development of socially produced regional substance in web-based life. Notwithstanding existing materials, for example, proverbs, nearby sayings, exhortation, and fables that are discovered spread on the web



Edit with WPS Office

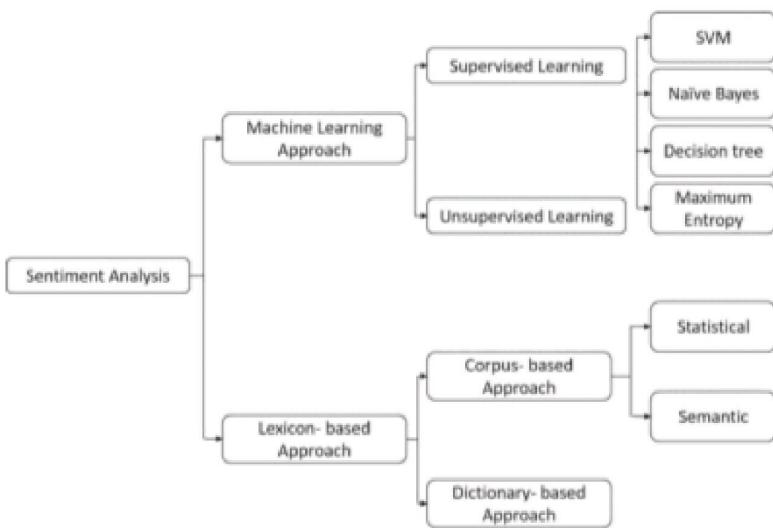


Fig. 1 The main techniques used in sentiment analysis

2 Related Work

In this section, we cover some important parts in the sentiment analysis as follows:

2.1 Sentiment Analysis and Sentiment Analysis Mining

An essential part of the information-gathering form has eternally been to find out what different people believe. With the increasing availability and demand of sentiment analysis-rich devices such as online survey sites and private blogs, unique events and difficulties occur as people now can, and do, actively use learning technologies to attempt out and get the sentiment analysis of others. The unexpected eruption of movement in the field of sentiment analysis mining and sentiment analysis, which dispenses with the computational method of sentiment analysis, feeling, and subject trinity in the document, has thus happened at least in part as a right answer to the surge of investment in new systems that distribute direct with ideas as a first-class object [11]. Web is the most critical wellspring of getting up considerations, overviews for an item, and surveys for an administration or movement. A Bulky measure of surveys are created day by day on the internet about online items and articles. For instance, numerous people share their comments, surveys, and emotions in their language using online life systems, for example, twitter, etc. The prerequisite for structuring for various lingos is extending, especially as blogging and small-scale blogging destinations are getting to be prominent [12]. Exhibited a complete review of the methodologies of examination and diagrams of significant gaps in the writing.

There is a method to analyse the feelings of humans, which is the sentiment analysis. The internet provides accessible textual data that is growing every day. Online websites for shopping give shoppers the possibility to record their notes and reviews about goods sold, and that has to improve products selling and to enhance the satisfaction of shoppers [8, 13–15]. Dealing with a huge amount of records or comments in sentiment analysis is hard because the process using this amount is difficult to extract the general sentiment analysis. This huge amount of sentiment analysis can't be analysed manually, hence the



Edit with WPS Office

automatic method of sentiment analysis has an important role in solving this problem.[16] proposed another weighing plan for content examination purposes, [2] used the weighing plan for content examination. Just as, [4] proposed awaiting approach combined with the term frequency and inverse document frequency (TF*IDF) for Arabic assessment arrangement on administrations" audits in Lebanon nation. Sur- vets are about open administrations, including lodgings, eateries, shops, and others. They collected the materials from Google and Zonate, which have come to 3916 surveys. Trials show three centre discoveries: (1) the classifier is sure when used to anticipate positive audits. (2) The model is one-sided on anticipating audits with a negative feeling. Finally, the low level of negative audits in the corpus adds to the modesty of the calculated relapse model. Other optimization techniques can be sued to solve this problem [17–25].

2.2 Sentiment Analysis for Health Care

In this work [26], the aspects of sentiment in the medical field and possible use-cases are reviewed. Through the research review, the state of the art in healthcare environments is summarized. To learn the linguistic peculiarities of thought in medical texts and to get open research issues of sentiment analysis in medicine, they make a quant- tentative evaluation concerning term language and sentiment combination of a dataset of clinical stories and medical group media obtained from six separate references. Sentiment analysis in healthcare deals with the healthcare problems of patients [7]. The sentiment analysis of patients is taken to solve their problems quickly and this helps the decision-makers to make plans and useful changes. Sentiment analysis is used in different fields. Healthcare analysis-based study displays the powerful points of medicines and services [1, 5].

2.3 Sentiment Lexicons for Health-Related Sentiment Analysis Mining

Natural language processing includes an approach called sentiment analysis mining: this approach recognizes the passionate tone behind a frame of the document. This is a popular way for organizations to determine and categorize sentiment analysis about a product, service or idea [27–34].For this task, there are some resources required like polarized lexicon. Sentiment analysis mining in healthcare is not explored well, partially because some confidence is provided to patients and their sentiment analysis, and there are many patients used social media. They are inspired in sentiment analysis mining of reviews about medications. Firstly, they define the origin of lexicon, including sentiment analysis words from the common domain and their polarity [35]. Then they perform the production of a medical sentiment analysis lexicon, based on a corpus of medication reviews. The Arabic language experiences the absence of accessible enormous datasets for AI and sentiment analysis applications. This work introduced a huge dataset called BRAD, which is the biggest book reviews in Arabic Dataset [36]. This dataset contains 490,587 inn surveys gathered from the Booking.com site. Each record contains the survey message in the Arabic language, the commentator's evaluating on a size of 1 to 10 stars, and different qualities about the lodging/analyst. They make accessible the full unequal dataset just as a decent subset. To analyse the datasets, six prevalent classifiers are utilized using Modern Standard Arabic (MSA). They test the slant analysers for extremity and rating groupings. Moreover, we actualize an extremity dictionary-based slant analyser. The discoveries affirm the viability of the classy fires and the datasets. Our centre commitment is to make this benchmark- dataset accessible and available to the examination network on the Arabic language. Information explanation is important to sort content into two classes, the first thin decent comment rule is needed, and the second is building up what is required to fit the bill for each class. In this paper [37], they present a novel way to deal with naturally developed corpus for Algerian tongue (A Maghrebi Arabic vernacular). The development of this corpus depends on an Algerian feeling vocabulary that is likewise built naturally. They



Edit with WPS Office

worked on two broadly utilized contents using Arabic internet-based life: Arabic and Arabism. The accomplished F1-score is up to 72% and 78% for an Arabic and Arabism test, respectively.

2.4 Sentiment Analysis on Health Services

Social media system appearance produces huge numbers of important data that is accessible online and accessible easily. Several users share discussions, photos, and videos, data, and viewpoints on various social media websites, within Twitter presence one of the common modern [9, 38, 39]. Consistently a lot of emotional data is created through informal organizations, for example, Facebook and Twitter. The abstract data suggests the suppositions, convictions, emotions, and frames of mind that individuals express towards various subjects of intrigue [40]. In addition, this sort of data is critical for organizations, associations or people, since it enables them to do activities that advantage them. In addition, assessment examination is the field that reviews emotional data through character-istic language handling, computational etymology, data recovery, and information mining strategies. Sentiment analysis is exceptionally helpful in different spaces, for example, legislative issues, highlighting, the travel industry, among others. In fact, the social insurance space suggests a huge region of the chance to acquire adman-tags utilizing sentiment analysis examination, for example, getting data about the patients' state of mind, illnesses, unfriendly medication responses, plagues, among others. Nevertheless, the social insurance space has been almost no investigated. It is very difficult to extract useful information from Twitter, because this data is unregulated, and this a big challenging responsibility. There are many Arab users on twitter, and these users are used the Arabic language to write their posts and tweets. In the English language, there are a lot of researches about sentiment analysis, but in the Arabic language, it is very poor. They introduced an Arabic language dataset about sentiment analysis on medical services received from Twitter [41].

2.5 Sentiment Analysis Techniques in Healthcare

A large number of subjective knowledge's is created through social media websites such as Facebook, Twitter. This knowledge indicates the ideas, views, emotions, and beliefs that people pass across various issues of matter. It has a revolutionary point for several organizations or people in order to provide information about what represented to the goods or services, which allows them to carry out activities that make advantages for them, such as how to make better choices, better promotion operations or business strategies, among others [42]. Individuals are uncommon to discuss about their medical issues with one another and, it is extremely poor to see about their practical wellbeing circumstance. Consequently just on Twitter, clients' made tweets made out of the news, governmental issues, life discussion which can likewise be connected for completing an assortment of examination purposes. Hence, the human services framework is created to help the experts to effectively check their conduct sentiment analysis depend on Twitter information. Most extreme Entropy classifier (MaxEnt) is utilized to perform sentiment analysis on their tweets to propose their wellbeing condition (great, reasonable, or terrible). It is communicating with Twitter information (huge information Condition) thus, the Internet of Things (Iota) based huge information preparing structure is worked to be effectively dealt with a lot of Twitter client' information [43]. The method of personal knowledge and the knowledge to recognize the feelings Sentiment

3 Future Work

analysis methods would produce aggregated personal decisions on the healthcare inquiries. A suggested way



Edit with WPS Office

can be made on top of such inquiry analysis that would suggest medications, procedures, experts in the region, important health care stations etc. based on the individual information presented by the different patients. More complex methods can be used in such methods to control spam. The use of science, fundamental knowledge and machine learning methods must merge to decrease the influence on decisions. More such parameters are to be determined, through which the efficacy of content can be checked. Producers may also be assigned based on the legality of the content they provide wher.

4 Challenges of the Natural Language Processing

The Natural Language Processing has several difficulties that can change the appear- ance of the sentiment analysis in many aspects [48]. Some of the certain difficulties are related to the kind of data while others are obvious to any type of analysing text. The current difficulties in Natural Language Processing can be divided as follows [35]:The document level difficulties are associated with the inquiry text that can have reviews, which are only found in blogs. Blog reports provide annotations that nor- mally become within a forum. These reports have feelings that are particular to the field. Based on personal therapy and diversity in natural languages, personals display themselves negatively. Sentiment analysis spamming is also a very sensitive point where somebody gives incorrect sentiment analysis prepared for serving or reducing special-purpose things (manually). There may also be displays as inspection reports that would become nothing to do with the destination actuality in the discussion [49]. Regrettably, there is somebody and organizations included in the market of seen- timing analysis study spamming. Specialists in sentiment analysis are demand as diverse as half of the inspections to be spammed on any popular display roots. Ute- libation of special sentiment indications are also required as the authors' content. Finally, grammar errors, local slangs are other generally handled difficulties.

5 Discussions and Overview

There are many challenges facing sentiment analysis techniques. For instance, the complexity in the way of the people to express sentiment analysis, lexical content in the text, irony, and implication. Therefore, various techniques have been used



Edit with WPS Office

Table 1 Shows the various techniques have been used in sentiment analysis.

Author(s)	Technique	Approach	Accuracy
[49]	Lexicon	Corpus- based Approach	54%
[43]	SVM Naïve Bayes Maximum Entropy	Supervised	
[50]	Max Entropy	Supervised	Positive (34.34%) Negative (22.18%) Neutral (43.46%)
[16]	SVM Naïve Bayes	Supervised	94.18%
[50]	Logistic Regression	Supervised	
[12]	SVM Decision tree Nearest Neighbor (k-NN)	Supervised	Below 50%
[1]	Lexicon	Corpus- based Approach	
[13]	Lexicon	Corpus- based Approach	
[37]	Lexicon	Corpus- based Approach	78%
[40]	SVM Decision tree	Supervised	
[46]	Convolutional Neural Networks (CNNs)	Supervised	91%

(see Table 1) to solve the sentiment analysis problems. However, each technique has advantages and disadvantages. For example, the Naïve Bayes is efficient and fast computation without influenced by irrelevant features. However, it assumes indecent- dent attributes. While Maximum Entropy does not suppose statistical independence of random elements [50]. However, it needs more efforts from the human in the form of additional resource [7]. Finally, the main advantage of the lexicon is that since the accuracy of the comments achieved by humans is not guarantee. Nevertheless, this technique consumes immense time [1].

6 Conclusion

Sentiment analysis is a necessary way to help people in getting a recommendation and read knowledge. This technique aims to analyse the social media, wherever a problem highlighted may only contact the necessary authorities if they notice it immediately. It is impossible through the social media and various user content to get the right recommendations. Sentiment analysis is automating this process. Sentiment analysis aim to get more information to assist users to get the right decision about the studied. The sentiment analysis methods are applied to data mining and machine learning to adjust this difficulty. Supervised procedures with high precision

can be employed to further sensible proposals for finding close bound problems. Unsupervised techniques are less costly and can be used to investigate big data. The performance of the sentiment analysis process is categorized with interest rates for the potential sources. It can be supported with graphical devices to be



Edit with WPS Office

more decisive to users. Reviews may additionally be applied to set feather the conclusions. This state is still greatly from being ready with new sub-streams known as disturbance analysis, expression analysis, preference analysis, risk analysis, etc. The demand and insufficiency of these decisions can be recognized from the evidence that it has previously been done for financial goods while it is still active as a research difficulty. Finally, the optimization way can be used to deal with this problem; it gave promising results in solving several problems.

References

1. L. M.T. Khan, S. Khalid, Sentiment analysis for health care, in Big Data: Concepts, Methodology
2. M. Shahab, A.T. Khaddar, M.A. Al-Beta, L.M. Aaliyah, Hybridizing cuckoo search algorithm with hill climbing for numerical optimization problems, in 2017 8th International Conference. on Information Technology (ICIT), IEEE (2017, May), pp. 36–43
3. L.M.Q. Aaliyah Feature selection and enhanced Krill Head algorithm for text document clustering, in Studies in Computational Intelligence (2019)
4. H. Malik, H. Haddad, C. Bacchic Ali, L. Bamboozle, Tunisian dialect sentiment analysis: natural language processing-based approach. Computation y Sistema's 22(4) (2018)
5. M.M. Mostafa, NR. Nebo, Sentiment analysis of Spanish weeds of Arabic origin related toilsome: a social network analysis. J. Lang. Teach. Res. 8(6), 1041–1049 (2017)
6. G. Vicodin, R.M. Chandrasekhar, Sentiment analysis and sentiment analysis mining: a survey, Int. J. 2(6), 282–292 (2012)



Edit with WPS Office