

A Survey on Sentiment Analysis Techniques in the Medical Domain

M Srisankar^{1*} and KP Lochanambal²

¹Research Scholar, Department of Computer Science, Government Arts College, Udumalpet, Bharathiar University, Tamilnadu, India

²Assistant Professor, Department of Computer Science, Government Arts College, Udumalpet, Bharathiar University, Tamilnadu, India

***Corresponding Author:** M Srisankar, Research Scholar, Department of Computer Science, Government Arts College, Udumalpet, Bharathiar University, Tamilnadu, India.

Received: December 23, 2023; **Published:** January 09, 2024

DOI: 10.55162/MCAES.06.157

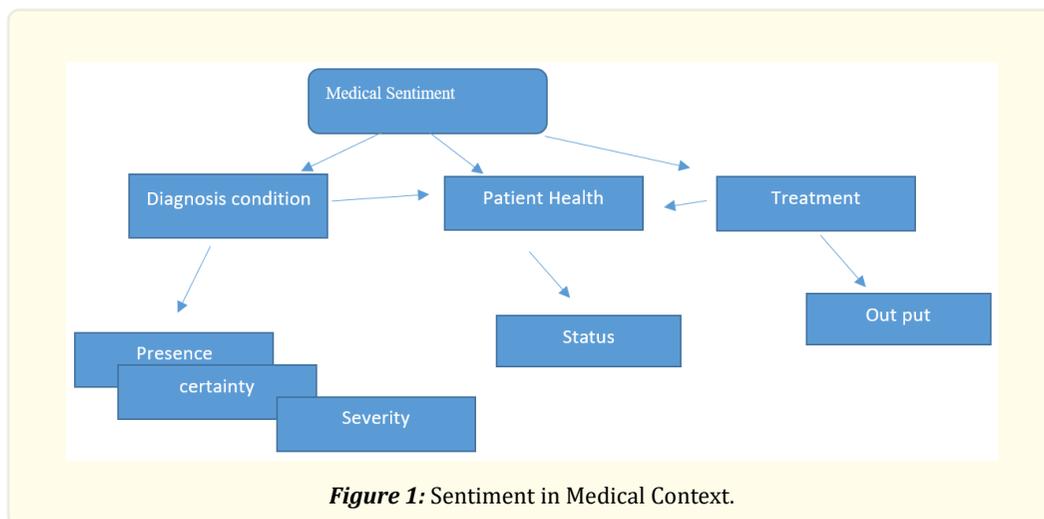
Abstract

Twitter is a microblogging platform that allows users to send and read 140-character messages, or “tweets.” A lot of unstructured, free content linked to healthcare is shared on Twitter, which is becoming a well-known platform for restorative research. One type of information mining that provides an assessment of the identification opinion investigation process in characteristic dialect preparation is estimation investigation. Computational phonetics is used to infer and analyse mental knowledge of the internet, social media, and related references through content analysis. The information gathered truly assesses the attitudes or mindsets of the global community towards specific goods, people, or issues and revealed the pertinent duality of the data. Sentiment analysis is applied in various reveal the contextual duality of knowledge and the attitudes of society towards particular products, persons, or ideas. Many industries, including the health care industry, use sentiment analysis. Social media and websites that rate medical issues are only two examples of the vast amounts of online resources for healthcare information that are not carefully reviewed. Using medical data to optimise patient outcomes and raise the standard of healthcare is only one of the numerous advantages that come with sentiment analysis. The techniques for sentiment analysis that are utilised in the medical area are the main topic of this review article.

Introduction

Sentiment analysis (SA), also referred to as opinion extraction, is a branch of natural language processing (NLP) that classifies sentiment in loose text mechanically. Its roots can be traced lower back to the Nineteen Nineties, with methods for viewpoints classification [1], prediction of adjectives semantic orientation [2], subjectivity classification [3], and so on. however, the appearance of net 2.0 and the ever-growing availability of consumer-generated records, including product and service evaluations, in addition to the proliferation of social networking platforms, have all contributed to its speedy expansion. SA has been observed to be beneficial in various societal settings, which includes enterprise, financial system, and authorities [4-8]. This observe specializes in health apps, which are described as “a situation of whole mental, physical, and social wholeness, rather than genuinely the absence of diseases or infirmities” [9]. properly-being is described as a perceived or subjective condition, this means that that it can vary notably among individuals in otherwise comparable occasions [10]. This reality makes properly-being an superior case study for SA. concerning fitness troubles, of direction, due to this, properly-being is an notable case have a look at for SA. Of path, when it comes to health troubles, modern-day society is targeted on poor occurrences like infection, damage, and incapacity [11], making SA hard in this subject. as an instance, having a good quality of lifestyles for a patient with a long-time period illness will depend less at the absence of accompanying *Corresponding creator: elmendili.saida@uit.ac.ma problems and extra on how nicely they are controlled and managed. The unfavourable

difficulty of fitness symptoms, however, tends to shift the SA's consequences towards the poor stop of the range. nowadays, each doctors and sufferers use on-line platforms such, social media, and web sites to specific their view on health troubles [12]. Informatics is described as "the technological know-how of information dissemination and transmission in a web medium, drastically the internet, or in a populace, with the overriding motive of influencing public health and public coverage" [13]. actual-time information collection and processing is feasible from social media together with Twitter, with the capacity to survey public opinion (sentiment) on a given subject matter [14]. Social media has been defined by using Bates and associates as a "perfect typhoon" for affected person- targeted healthcare, which gives a treasured source of data for the public and healthcare institutions [15]. Twitter is a part of this, because it is straightforward to use, cheap, and accessible. There are 955 million followers on Twitter nowadays, that is a cell microblogging and social networking carrier. There are definitely 955 million registered customers on Twitter who can share submit containing textual content, movies, snap shots or hyperlinks to external assets. absolutely one-1/3 of humans with social media profiles use Twitter, and 75% of them access it from a cellular tool to explicit an opinion [16]. Sentiment analysis examines the content of natural language in unfastened textual content, i.e., the individual phrases and signs utilized in a text message, to determine the intensity of advantageous and poor critiques and emotions. Sentiment analysis the usage of social media is a widely studied subject matter already [17]. they're commonly based on textual content classifiers or device learning based processes. They tend to be enterprise oriented, high priced, and consciousness on capturing evaluations about a particular chosen product or branch [18]. In this studies, they provided a qualitative and specific measure of sentiment the use of content classifiers together with "humour" or "sarcasm." appropriately studying sentiment in an automated way is hard because of the subjectivity, flexibility, and creativity of the language used [19]. The reason of this observe was to study the techniques used to measure sentiment for Twitter-primarily based healthcare research. the first objective become to take a look at which techniques of sentiment evaluation had been used and in what healthcare context.



The motive of this study become to examine the techniques used to degree sentiment for Twitter-primarily based healthcare research. The primary objective became to have a look at which strategies of sentiment evaluation had been used and in what healthcare context. the second goal changed into to take a look at the quantity to which the strategies had been practiced and verified for the look at statistics, and whether a rationale for his or her use become furnished have been created in the subject of net science and internet surfing. This discipline is also known as sentiment analysis or opinion extraction. In medicine and health care, sentiment evaluation is of developing hobby.

Related Work

Methods for figuring out and reading subjectivity, as well as opinions and sentiments offered in textual form, created in the area

of internet technological know-how and web browsing. This discipline is likewise referred to as sentiment evaluation or opinion extraction. In remedy and health care, Sentiment evaluation is of developing interest. Sentiment analysis research began in 2004 [20] with the investigation of customer comments and online news for opinions. The challenge at hand was to discern between positive and negative viewpoints, as well as subjective and objective sections of a document. Later on, additional tasks were added: emotion analysis determines the emotional category of texts [21], whereas intensity focuses on detecting the many levels of policies or feelings (e.g., very positive, very sad). Popular assessment Sentiment analysis studies began in 2004 [22]. With the investigation of client feedback and on line news for evaluations. The venture to hand turned into to parent between high-quality and bad viewpoints, as well as subjective and objective sections of a record. Later on, additional tasks have been delivered: emotion analysis determines the emotional category of texts with the investigation of customer whereas intensity focuses on detecting the many levels of policies or feelings (e.g., very positive, very sad). The classification problem is addressed in existing sentiment analysis systems built for processing unstructured text in Web media: a classifier is trained to determine the level of polarity at the sentence or document level. Pang et al., for example, proposed a number of supervised methods, including support vector machines (SVM), naive Bayes classification, and maximum entropy [23]. A solution for recognizing the text semantic focus has been developed by Turney, who suggested an unsupervised approach. The principle precept is that the sensation is expressed through opinion statements and is expressly referred to inside the text. Adjectives, adverbs, and particular nouns, on the other hand, normally suggest sentiment in free writing. Therefore, traditional sentiment analysis is based totally at the evaluation and detection of those opinion words (semantic capabilities). the usage of a lexical survey, those are then analysed with the aid of sentiment analysis algorithms to find opinion terms and their rules in a report. in addition to lexicon research to find opinion phrases, other capabilities can be extracted from texts to be processed with the aid of sentiment analysis techniques. Those consist of lexical attributes along with unigrams, bigrams, and phrases in elements of speech.

Sentiment lexicons

A number of the methods used in the sentiment analysis are actually primarily based on the usage of sentiment vocabularies, which provide an outstanding basis for the analysis of sentiment in the reputation of words and terms in a language. They are an critical basis for the identity of expressions and speech patterns of sentiment in herbal language texts. a few examples of current lexicons available include SentiWordNet (SWN) [33], WordNetAffect, fashionable Inquirer, and the Sentiment Lexicon (SL). They both have words in them and assign sentiment rankings or categories to them. Lexicographers can create lexicons through hand or by studying corpora. They often ignore the nuances and differing meanings of terminology whilst hired throughout domain names, or they fail to offer express facts approximately other topics. one of the most broadly used sentiment lexicons is SentiWordNet. It attributes to each WordNet synset 3 emotions ratings: positivity, negativity, and objectivity. SentiWordNet has a ramification of word definitions. however, it is not laid out in which area a phrase has a special meaning (for example, if right is honestly a navigational time period or a term which means "correct"). word NetAffect assigns one or extra affective labels to a group of WordNet synsets that constitute emotional feedback. Mohammad and Turney have built EmoLex, a manually created emotion lexicon the usage of Amazon's Mechanical Turk service. The Lexicon of Subjectivity by way of Wilson et al [35] includes 8221 precise term expressions of subjectivity mixed with their related polarity. to be able to use these global emotions assets, subject-particular vocabulary could be in addition advanced based on the phrase's meanings area specificity. The most commonplace domain words, along side their highstage semantic linkages and event patterns, could be protected in a multidimensional 86f68e4d402306ad3cd330d005134dac area lexicon. enhancing an present vocabulary is the maximum direct way to cover the area-specific environment. Combining a widespread lexicon with a domain-particular lexicon is one way to improve the lexicon's insurance. Whilst traditional methods have in most cases focused on locating the specific expression of sentiment, furnished an annotation method that took into consideration each active useful and energetic masculine occurrences across opinion entities, resulting in an extension of the Subjectivity Lexicon. mainly, four forms of "top for" and "bad for" of events have been advanced: destruction, generation, benefit or loss, advantage or injury. occasions are expressed as a text triplet: agent, gbf, object, which indicated that a topic (noun) including someone or employer had an amazing or adverse impact on the object (noun). The author's cautioned viewpoints are indexed inside the mixed phrases from Senti Word Net and Subjectivity Lexicon preferred-motive lexicons. The merged lexicons had been then supplemented with opinion words from medical critiques. even as

extending, the differences in terms of polarity between the general and scientific fields were tested. The authors observed that during pharmacological publications, some terms that are generally thought to be impartial are honestly bearers of opinion. sooner or later, the usage of a fundamental scoring gadget, it changed into showed that using the mixed vocabulary yielded better effects than using extensive lexicons of sentiment. Sentitivity Lexicon, trendy Inquirer, SentiWordNet, and Moby have been used to assess the accuracy of feeling reputation. textual content is from six classes of preferred domain datasets, which includes inn reviews, e-book reviews, films and conversations, changed into delivered to the lexicons. The findings discovered that the extent of accuracy of individual Lexicons varied by way of area. SentiWordNet, in instance, had the satisfactory accuracy (65-seventy one%) in four of the six domain names, whilst the subjective lexicon had the best performance (63-sixty five%) inside the closing. specifically, SentiWordNet finished the pleasant accuracy (65-71%) in four of the six fields, and the Subjectivity word list performed quality (sixty three-sixty five%) within the different two domain names. SentiWordNet's tremendous vocabulary insurance changed into one component.

Sentiment analysis

In a scientific setting In terms of familiar techniques for reading attitudes, the following sections move over the various tactics in further depth. Inside the discipline of drugs, sentiment evaluation is labeled according to the textual source (e.g., clinical internet content material, biomedical publications, scientific notes), the technique (e.g., based totally on polarity, classification, rules, machine learning), and type of analysis (final results type), and degree (e.g., phrase level, sentence degree). The maximum critical methods are unique inside the phase that follows. In current research, feeling is regularly taken into consideration as a polarity, i.e. a wonderful, negative or neutral polarity toward a given subject. this type of categorization is relevant while. This categorization is applicable when, in texts, opinions are expressed approximately someone (e.g., a medical doctor), a drug, or a scientific device. however, the sentiment can be even extra essential. rather than humans or merchandise, in which emotions primarily encompass appreciation or dislike of a person or good, views or feelings about medications, treatments, or even diagnoses have even more faces and are phrased in many different words: - Opinions about medical doctors, medical devices, and medications, - Personal sentiments about one's own or others' health conditions, -Complications that have occurred, - Facts and experiences about a particular treatment, diagnosis or medication. A treatment can be painful, but useful. Complications may have occurred, but the treatment was still successful. A diagnosis can be frightening, but not really serious or fatal. A medication can have serious side effects. How can we characterize feeling in medical texts? It is clear that sentiment or opinion in medical social media can be expressed differently than sentiment in news or product reviews. A sentiment can also be described by a symptom that reflects a person's health condition. It is not just a feeling, it is characterized by symptoms, by pathological terms. Since in social media, a mixture of facts and experiences is expressed, it may be necessary to distinguish factual information from experiences. Most sentiment analysis work in the field of medicine is focused on Web data such as medical blogs or discussion groups, with the goal of mining or studying the opinions of individual patients or to measure the quality of healthcare. As an example, an approach was implemented to split evidence-based texts from evidence-based texts In order to evaluate the accuracy and validity of the patient generated content. With the assumption that content that is factual is actually better than affective content, as it is more informative (as opposed to moods and experience). With the concept that factual material is better than emotive material since it provides more information, a system was created to evaluate factual material in medical social media utilizing subjective terms and a medical ontology. Existing approaches to sentiment analysis from medical online data are typically machine learning-based or rule-based, similar to general sentiment analysis. The majority of research has focused on polarity ranking: Xia et al. provided a multi-step technique to user opinion classification. Their proposed approach identifies the subject and the polarity being expressed towards that subject. An Fmeasure of about 0.67 was also reported. For the classification of sentiment in tweets, used a variety of classifiers, including Naive Bayes, decision trees, and support vector machines (SVMs). The texts were regarded as a jumble of words. Two clustering tasks were considered: three classes (positive, negative, and neutral) and two groups (positive, negative, and neutral) (positive, negative). An SVM classifier had the best Fmeasure of 0.69. The purpose of the work. was to find the polarity of the expressed sentiments of patients in online health resources communities. In particular, they used the domain dependent and domain independent sentiments characteristics as the two complementary points of view of a message and used them for message classification in a semi-supervised setting using a co-training algorithm to complete a classification of user message sentiments in an

online cancer support service community. This effort was later improved with features that are derived from a dynamic sentiment vocabulary, while the earlier work employed a more general sentiment lexicon to abstract the patterns. An additional perspective of the sentiment in patient feedback, namely speech features like expressivity and persuasiveness. A classifier was tested on a selected corpus of patients' feedback from NHS Choices. The outcomes show that the Naïve Bayes multinomial classification with features based on frequency can obtain the best accuracy (83.53%). In addition, the performance results demonstrated that a classifier model that was trained only on an expression corpus can be implemented directly on the persuasive corpus and achieve performance comparable to that of the corpus-based training with the same speech feature.

Sentiment Analysis for Medical Future

In conclusion, the study of medical texts (medical social media texts, clinical accounts) with regards to the analysis of sentiment includes the possible need to meet the following research challenges: - Modeling the implicit clinical context and identification of the implicit sentiment, - Construction of a field-specific vocabulary of sentiment, - Context-specific determination of sentiment, and - Modeling various aspects of the condition of the patient. Some other aspects are the identification of the view carrier and the incorporation of time. An operation may start usually, but may become crucial. Because of this, the health feeling must also be regarded in time. Time is supplied by the document timestamp or sometimes in the documents themselves, or clinical data can be sorted into phases of treatment. As other research has already shown, negative statements are used very much in clinical narratives. In the sentiment analysis, it is very important to detect the use of negations because the polarity can be switched (e.g., in no complaint of pain when asked).

Conclusions

This article presents the various ways in which the subject of sentiment is used in the field of medicine and identifies areas for future study in the areas of medical sentiment analysis. While the existing work analysis of sentiment from medicinal texts has focused on medical social media and biomedical literature, future work should also focus on the analysis of clinical documents. This analysis could be used for numerous purposes. Uncertainty, attitudes, and implicit emotions might all be collected and used into therapeutic decision-making development of a domain-specific sentiment lexicon and techniques to judging sentiment based on the nature of the content could be the initial steps toward solutions for sentiment analysis.

References

1. Wiebe J and Bruce R. "Probabilistic classifiers for tracking point of view". *Progress in communication sciences* (1995): 125-142.
2. Hatzivassiloglou V and McKeown KR. "Predicting the Semantic Orientation of Adjectives". In: *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*. 1997 Presented at: ACL'98/EACL'98; Madrid, Spain (1997): 174-181.
3. Wiebe JM, Bruce RF and O'Hara TP. "Development and Use of a Gold-standard Data Set for Subjectivity Classifications". In: *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*. 1999 Presented at: ACL'99; College Park, Maryland, USA (1999): 246-253.
4. Hu M and Liu B. "Mining Opinion Features in Customer Reviews". In: *Proceedings of the 19th national conference on Artificial intelligence*. 2004 Presented at: AAAI'04; San Jose, California, USA (2004): 755-760.
5. Hu M and Liu B. "Mining and Summarizing Customer Reviews". In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004 Presented at: KDD'04; Seattle, Washington, USA (2004): 168-177.
6. Bollen J, Mao H and Zeng X. "Twitter mood predicts the stock market". *J Comput Sci* 2.1 (2011): 1-8.
7. Efron M. "Cultural orientation: Classifying subjective documents by cociation analysis". In: *Proceedings of the AAAI Fall Symposium on Style and Meaning in Language, Art, and Music*. 2004 Presented at: AAAI'04; San Jose, California (2004): 41-48.
8. Ramteke J., et al. "Election Result Prediction Using Twitter Sentiment Analysis". In: *Proceedings of the 2016 International Conference on Inventive Computation Technologies*. 2016 Presented at: ICICT'16; Coimbatore, India (2016): 1-5.
9. World Health Organisation. Geneva, Switzerland: World Health Organisation; Constitution of the World Health Organisation

- (2006).
10. Huber M., et al. "How should we define health?". *Br Med J* 343 (2011): d4163.
 11. Berg O. "Health and quality of life". *Acta Sociologica* 18.1 (1975): 3-22.
 12. Afyouni S, Fetit AE and Arvanitis TN. "#DigitalHealth: exploring users' perspectives through social media analysis". *Stud Health Technol Inform* 213 (2015): 243-246.
 13. Eysenbach G. "Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet". *J Med Internet Res* 11.1 (2009): e11.
 14. Eysenbach G. "Infodemiology and infoveillance tracking online health information and cyberbehavior for public health". *Am J Prev Med* 5.Suppl 2 (2011): S154-S158.
 15. Rozenblum R and Bates D. "Patient-centred healthcare, social media and the internet: the perfect storm?". *BMJ Qual Saf* 22.3 (2013): 183-186.
 16. Ofcom. The communications market report (2015).
 17. Lunden I Techcrunch. Mobile twitterm+ (75%) access from handheld devices monthly, 65% of ad sales come from mobile (2013).
 18. Pang B and Lee L. "Opinion mining and sentiment analysis". *Foundations and trends in information retrieval* (2008): 1-35.
 19. Liu B and Zhang L. "A survey of opinion mining and sentiment analysis". *Mining Text Data* (2012): 415-463.
 20. Nasukawa T. "Sentiment analysis: capturing favorability using natural language processing". 2003 Jan 01 Presented at: Proceedings of the 2nd International Conference on Knowledge; Sanibel Island, FL, USA (2003).
 21. Chew C and Eysenbach G. "Pandemics in the age of twitter: content analysis of Tweets during the 2009 H1N1 outbreak". *PLoS One* 5.11 (2010): e14118.
 22. Mohammad S. "9 - Sentiment analysis: detecting valence, emotions, and other affectual states from text". *Emotion Measurement* (2016): 201-237.
 23. Hu M and Liu B. "Mining and summarizing customer reviews". In: *Proceedings of the 10th ACM SIGKDD international conference on knowledge discovery and data mining, KDD '04*. New York, NY, USA: ACM (2004): 168-77.
 24. Mishne G. "Experiments with mood classification in blog posts". In: *1st Workshop on stylistic analysis of text for information access* (2005).
 25. Baldoni M., et al. "From tags to emotions: ontology-driven sentiment analysis in the social semantic web". *Intell Artif* 6.1 (2012): 41-54.
 26. Pang B, Lee L and Vaithyanathan S. "Thumbs up?: sentiment classification using machine learning techniques". In: *Proceedings of the ACL-02 conference on empirical methods in natural language processing - Volume 10, EMNLP '02*. Stroudsburg, PA, USA: Association for Computational Linguistics (2002): 79-86.
 27. Pang B and Lee L. "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts". In: *Proceedings of the 42nd annual meeting on association for computational linguistics, ACL '04*. Stroudsburg, PA, USA: Association for Computational Linguistics (2004).
 28. Turney PD. "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews". In: *Proceedings of the 40th annual meeting on association for computational linguistics, ACL '02*. Stroudsburg, PA, USA: Association for Computational Linguistics (2002): 417-24.
 29. Pang B, Lee L and Vaithyanathan S. "Thumbs up?: sentiment classification using machine learning techniques". In: *Proceedings of the ACL-02 conference on empirical methods in natural language processing - Volume 10, EMNLP '02*. Stroudsburg, PA, USA: Association for Computational Linguistics (2002): 79-86.
 30. Pontiki M., et al. "Semeval-2014 task 4: aspect based sentiment analysis". In: *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)* (2014): 27-35.

Volume 6 Issue 2 February 2024

© All rights are reserved by M Srisankar., et al.