Acta Univ. Sapientiae Informatica **15**, 2 (2023) 330–344

DOI: 10.2478/ausi-2023-0020

Enhancing healthcare services recommendation through sentiment analysis

Muhammad Rizwan Rashid RANA
University Institute of Information Technology, Pir Mehr Ali Shah Arid Agriculture University
Rawalpindi, Pakistan
email: rizwanrana315@gmail.com

Asif NAWAZ
University Institute of Information Technology, Pir Mehr Ali Shah Arid Agriculture University
Rawalpindi, Pakistan
email: asif.nawaz@uaar.edu.pk

Tariq ALI
University Institute of Information Technology, Pir Mehr Ali Shah Arid Agriculture University
Rawalpindi, Pakistan
email: tariq.ali@uaar.edu.pk

Ghulam MUSTAFA
University Institute of Information Technology, Pir Mehr Ali Shah Arid Agriculture University
Rawalpindi, Pakistan
email: gmustafa@uaar.edu.pk

Abstract. As technology advances, most people use social media sites like Twitter, Facebook, and Flickr to share information and communicate with others. The volume of free-text data is growing daily due to the widespread use of these social media platforms. These platforms contain a substantial amount of unstructured information. Patient opinions expressed on social media platforms play a significant role in healthcare improvement and impact health-related policymaking. In this research, we introduce a machine learning approach for the optimal identification of healthcare-related features. This approach is based on a novel synthetic

Key words and phrases: sentiment Analysis; opinion mining; entropy; feature extraction; quality of services;

method. Additionally, we employ an entropy-based technique to classify free-text comments from hospital data into positive, negative or neutral. The experimental results and evaluations show 85%, 82.3%, 78.2% and 87% accuracy between ratings of health care. We observed that there is a minor association between our technique, expert opinion and patient interviews. Through the use of machine learning techniques, we achieve an accuracy level that suggests we are capable of providing an accurate and reasonable assessment of the ideal healthcare center for a patient. Our proposed novel framework predicts the healthcare experience at hospitals based on patient reviews posted on social media. This innovative approach outperforms traditional methods, such as surveys and expert opinions.

1 Introduction

Patient feedback helps in improving overall QoS in healthcare systems [9, 11]. Conventional procedures of patient feedback through surveys and reports reveal sustainable development in health services. These techniques are expensive and require some basic questions that are conducted frequently. In the present era, patients around the globe share their healthcare experience on the internet, social media websites or demonstrate in health reports in the form of blogs on different online healthcare communities [10, 6]. However, such type of information is much unstructured and difficult to understand.

In the case of unstructured data, it is very difficult to understand about patient's experience. Studies demonstrate that 85% of grown-ups utilize the web, 26% of individuals perused another person's encounters about health and 12% of individuals utilize online reviews of hospitals or some therapeutic minding sites [5]. In recent years, sentiment analysis has turned out to be progressively prevalent for preparing internet-based life information on online networks, wikis, micro-blogging stages, and other online cooperative media [23]. Intended to characterize text use sentiment analysis which is a part of full effective computing research. But it can also classify sound and video into positive, negative or neutral [19]. A large portion of the writing is in the English dialect; however, the number of publications suffers from multi-lingual issues [12].

It is very crucial to understand the main attributes and behaviors of approaches such as opinion mining, sentiment analysis and natural language processing. For example, in elections forecast sentiment analysis classifies natural language data into different positive or negative emotions [2]. On the other hand, if we apply the same approach to health services, for the translation of literary data about the patient's experience on a marvelous scale use different analytical techniques [14]. As a result of its composition nature, it maintained a strategic distance from the investigative spotlight of ordinary quantitative analysis. Alemi et al., proposed a technique related to the use of sentiment analysis of patient reviews in the form of real-time patient surveys [1]. They show that the comments of a specific patient are either positive or negative, as they set different classes for those sentiments. Moreover, they suggest that we need to compare sentiment analysis with old methods to know about patients' experiences.

SODA and RedMat allow different patients to define their experience with their health and different services at all hospitals in a particular country. People add around a million reviews per day about different services, particularly health care, QoS and management services. It contains an average of about 700 reviews about various hospitals [13, 7]. These comments contain both rating and free-text descriptions. They also calculate the experience of patients with the help of a survey about hospitals.

The rest of the paper is organized as follows: section 2 intro machine learning for patient comments. Section 3 demonstrates the conceptual model which is purposed by us. Section 4 represents the experimental results. Discussion is delivered in Section 5. The conclusion and future work is represented in Section 6.

2 Machine learning for patient comments

We conducted a method test using patient feedback obtained from RedMet and the SODA Choices website. Our primary objective was to estimate patient responses based on their comments. To achieve this, we utilized a machine learning algorithm to categorize the comments into distinct groups. We then compared our results to manually assessed comments by domain experts and conducted interviews to verify accuracy.In our effort to recommend improvements in Quality of Service (QoS), we thoroughly investigated free-text patient comments across various categories, including General Medical Services (GMS), Health Services (HS), Social Services (SS), and Management Services (MS). We developed a predictive model to assess a patient's likelihood of choosing a particular hospital based on the feedback we received. Additionally, our model assessed the hospital's hygiene and patient treatment.

General practitioners and physicians provide essential services through GMS, which form the foundation of healthcare services. The hospital offers HS, encompassing emergency services and ensuring patient comfort during painful situations. MS refers to services integrated by the hospital administration to enhance the well-being of patients and improve their service ratings. SS encompasses services provided by hospital staff, including nursing and housekeeping, aimed at creating a healthy and conducive environment for a swift recovery. All of these services collectively contribute to the overall enhancement of hospital QoS.

We used a set of data to test the predicted accuracy of the process. After the very first step i.e., pre-processing, we collect 71% valid data from total comments. We also compare our predicted results with the patient's rating which is based on interview and expert ratings.

3 Proposed model

Keeping in mind the end goal to get more fruitful results, we ensured that the execution of our recommended technique would be comparable to or better than the currently accepted solutions to healthcare difficulties in the field of sentiment analysis, the feature identification module demonstrated that the proposed approach outperforms existing techniques on supplied data. The proposed model is shown in Figure 1.

3.1 Datasets

Most datasets were obtained from Medicare under terms and conditions of use and security caveats. We choose one of them from the Socrata [4], Open Data API (SODA) [22]. Data facilitated in Socrata destinations are accessed using SODA software. We can look for information in various categories, including healing facilities, nursing homes, hospices, long-term care, and supplier directories. The official Centers for Medicare and Medicaid Services information is available on this site as a resource (CMS). The SODA API supports the JSON format, the most commonly used configuration for API answers. Socrata recommends this format since it is the simplest and most productive. Clients can submit and follow up on reviews of medical and dental experts, psychologists, urgent care centers, group practitioners, and hospitals on RateMDs.com. Both patients and doctors will benefit from this site. Table 3.1 shows a description of the dataset.

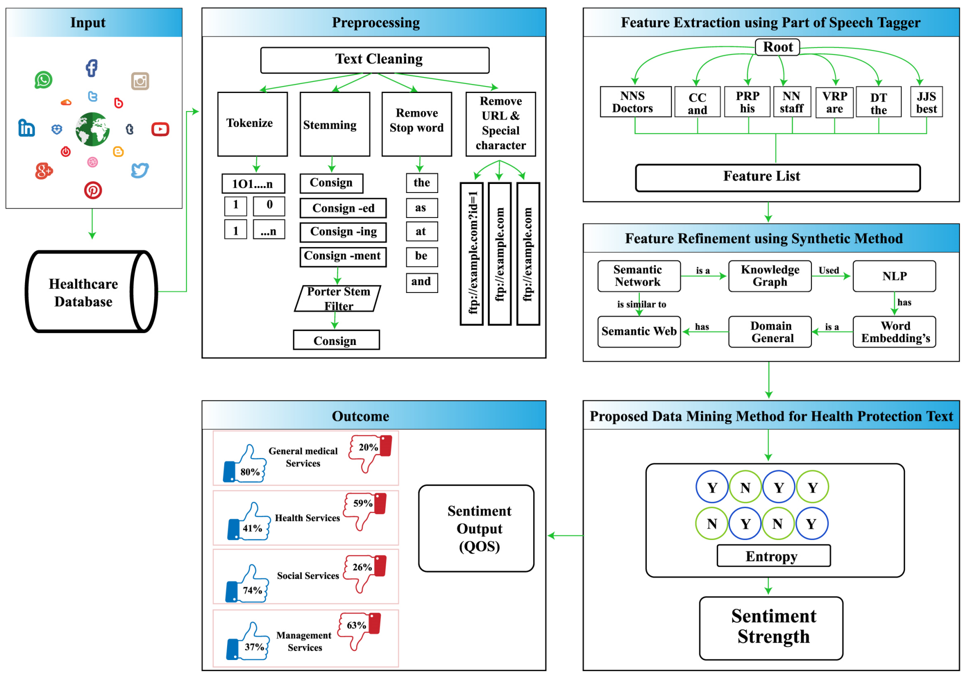

Figure 1: Proposed model

3.2 Pre-processing

The input text is split into sentences in this segment. The POS labeling and stemming are achieved with the help of Stanford CoreNLP [17, 3]. Words with positive polarity are used in sentences like "this is anything but a decent doctor's facility," however the refuting term NOT changes the polarity of a sentence. Furthermore, unigram characteristics do not show relationships between words in the material. Before extracting the unigram features, the negative word's impact should be reflected. Figure 2 depicts the data preprocessing.

3.3 Health protection feature extraction

Any item's reviews or tweets may contain unique characteristics that can be noted separately. Each sentence of a specific review associated with health care can be considered a sack of words at this level. The Modified POS tagger recognizes all highlights as well as opinions from the pack of words. A POS

#	Type	Dataset	Reviews/Tweets
1	RateMDs	Reviews of Patient	689
2	HOADF	Reviews of Physician	122,716
3	Medicare	Hospital	61358
4	Medicare	Nursing Home	40976
5	Medicare	Hospice	45787
6	Medicare	Long Term Care	35000
7	Medicare	Supplier Directory	20000

Table 1: Dataset breakdown [16]

tagging tool looks for grammatical form as well as linguistic linkages in other sentences in a sentence. Each word in a sentence is labeled as an action verb (VB), a noun (NN) and noun phrase (NNP), a proper noun (NNP), adjectives (JJ), and so on.

3.4 Health protection feature refinement using net (synthetic method)

Researchers have successfully applied machine learning algorithms to split sentiments in a document[15, 20]. However, as the feature set of data grows larger, the temporal complexity of these strategies grows. Furthermore, insignificant and repeated features play a role in determining the sentiment of a given document, causing the algorithm's accuracy to vary[8]. The primary goal of this step is to reduce the dimensionality of the feature space to obtain the most perfect feature and reduce computing costs

ConceptNet is a large-scale system that was launched in 1999 to understand the semantics of words [18]. Below are a few stages associated with the arrangement of our conceptual mode.

- ConceptNet, in all of its versions, includes social learning of the English dialect, and its sibling effort provides information on other well-known dialects.
- ConceptNet makes use of DBpedia's subset. It uses Wiktionary, a multilingual vocabulary, to extract knowledge from Wikipedia articles. This dictionary contains information on various topics, including health care.

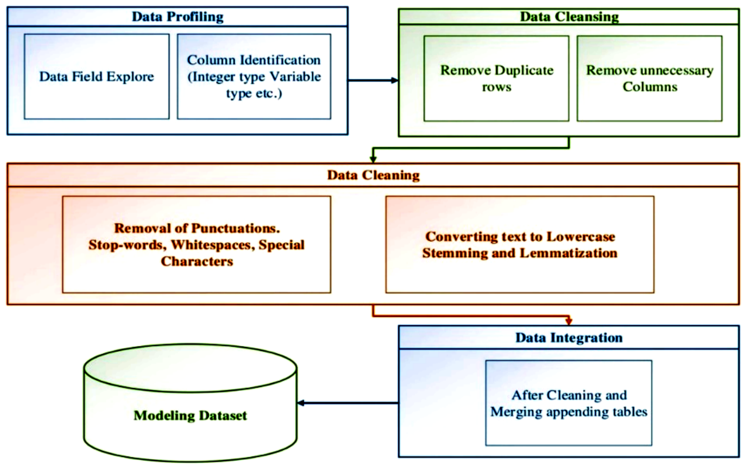

Figure 2: Proposed model

- For expanded information, the WordNet multi-dialect dictionary is also used.

- "Games with Purpose" provided some insight into people's feelings. The Japanese made this game for the GWAP challenge.

Figure 3 depicts a Concept Net improved model that was used in our suggested methodology as a tree with various healthcare-related properties.

3.5 Proposed data mining method for health protection text

The maximum entropy model is a modestly developed statistic model that was originally used to handle enormous amounts of authentic text. Gradually, it became clear that the maximum entropy model is also applicable to natural language processing [21]. The central concept is to provide a known event set, identify potential requirements based on it, and then choose a model that meets the imperative condition. Meanwhile, the probability distribution of the unknowns, distributed equally as may be expected given the conditions, is not completely understood. The probability is calculated using the maximum entropy approach [24].

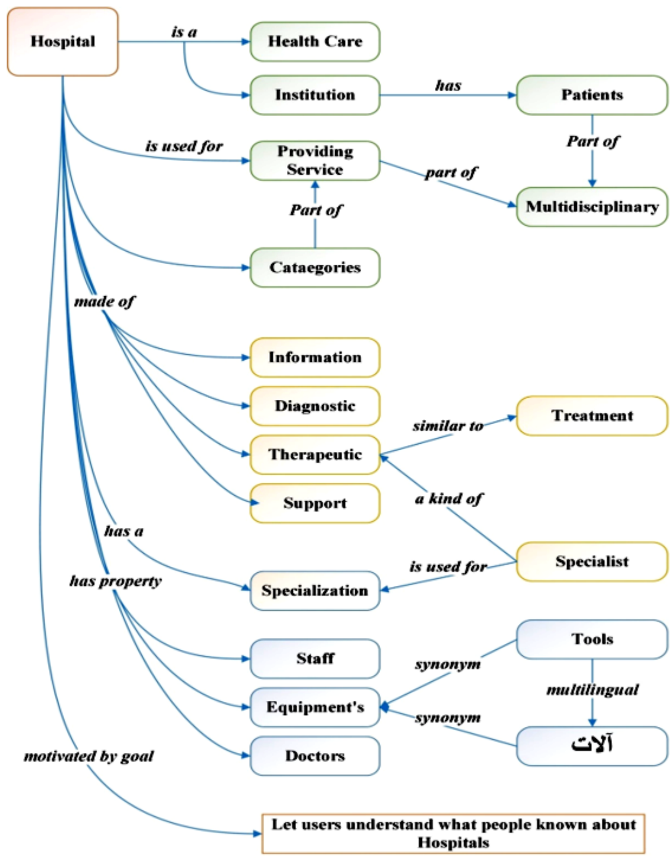

Figure 3: ConceptNet model for healthcare feature extraction

3.6 The text classification process

The preparation and testing of a classifier is required for text classification. The training data must preprocess and depict the text categorization process using the maximum entropy model. Different text features are created by handling training text following word segmentation, removing stop words, feature extraction, and word recurrence measurements. Portraying distinctive content as per the element of the greatest entropy work technique to ascertain the different parameters requires a maximum entropy classifier.

3.7 Strengths and limitations

Sentiment analysis using the technique is as good as the learning data set that we give as input. We can utilize numerous numbers of appraisals in the learning set than in different investigations. Also, with the use of sentiment analysis in medical services information, analysts need to prepare the framework themselves by evaluating reviews and attributing qualities to enable the technique to learn. We utilized a huge dataset that allowed us to specifically look at free text reviews and ratings posted by similar patients which help us to remove the potential base of reviewers during the assignment of a review.

Online comments left on a website without being solicited are likely to tend to gravitate toward models of both negative and excellent comments. Additionally, these online reviews are for the most part contributed by youthful young affluent people. Also, there are parts of patients' reviews that are difficult to process. There are a few words, for example, 'Irony', 'humor' and 'sarcasm' are regularly used in the English dialect and these words are hard for an algorithm to preprocess. The usage of earlier polarity enhanced the outcomes and gave some great results, yet, there were difficulties in understanding the context. Content that is trimmed again and again, for instance, "stank of pee" or "like a holy messenger", is effortlessly categorized as negative or positive. The sense of other ordinarily utilized expressions was hard to develop without knowing about their context. It was extremely hard to predict expressions without knowing the main context.

Currently, our approach cannot use such types of sentences that are looked clear on case to case basis. They are not using the most cutting-edge machine learning algorithms or approaches to classification selection in this early exploratory work, as observed in other industries [2, 24]. However, we believe that further work may have the capacity to embrace this.

4 Results

We had our proposed system assessed by a panel of experts from an NGO called 'X' that provides healthcare solutions. Two Medical Specialists, one Analyst, and two Quality Affirmation Specialists make up their experienced team. We've presented our proposed structure, as well as the knowledge base and user interface we've created. To their respectful expert group, our demonstration covers the strategy of research, organization and administration. We investigated whether such a medicinal services system on a specific topic has been built using our intended healthcare structure and whether it will aid in improving QoS in hospitals by removing ambivalence and collisions between patients and doctors. For the progress of QoS, we presented parameters based on GMS, MS, HS, and SS. They've started a rating system that goes from 0 to 10. Not Agreed is 0-3 points, moderately Agreed is 4-6 points, and Agreed is 7-10 points. We liquidated the outcomes based on agreement, partially agreed, and not agreed by applying the mean to all of the parameter's resultant predictions. The principle of expert opinions is based on three variables: agreed, disagreed, and somewhat agreed. Following the advice of the experts, it was widely agreed that the proposed technique is ideal for dealing with healthcare facilities in similar hospitals. The proposed approach to collecting patient data is simple to implement and aids in the reduction of comments and reviews.

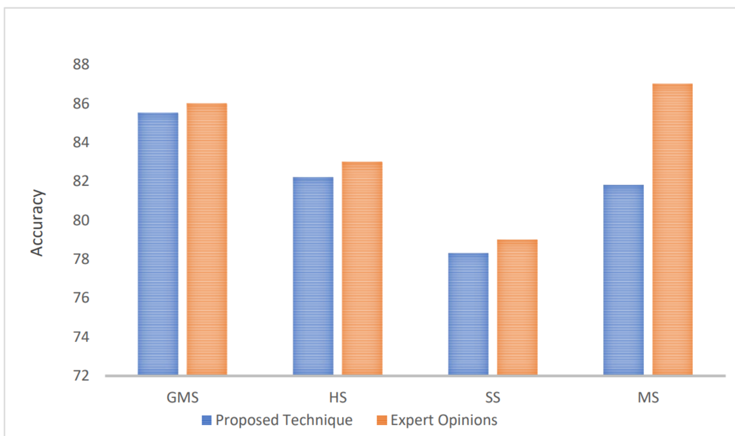

Figure 4: Comparison of proposed technique with expert opinion

Following the evaluation of our predicted method, we compare the achieved results to the interviews conducted with various patients who came to the hospitals at random. This is a poorly controlled process that is used on haphazardly selected patients who are admitted to hospitals and are suffering from various diseases. This interview policy applied to nearly 30 hospitals and included both general and specific questions. In these meetings, we simply select specific information, such as our anticipated topic. Every question has a set of character choices that range from "excellent" to "poor." The suggested analysis ranking differs from the patient ranking. In the first experiment, the obtained results are distinguished by the inclusion of expert suggestions. It demonstrates that the accuracy depicted as a completion measure of the suggested approach is optimistic. Each piece of hospital advice is a perfect match for expert opinions. Figure 4 depicts the graphical representation. The comparison of proposed technique with the interview-based results is shown in Figure 5. Figure 6 represents an overall comparison of the suggested technique with expert opinions and interviews, demonstrating that it performed well in terms of performance measurement. This comparison clearly shows that, given the complexity of the free text and the difficulties of obtaining expert opinions and conducting interviews, the proposed technique performs admirably. Without the challenges of conducting interviews and appointing an expert panel, the accuracy gained is nearly the same.

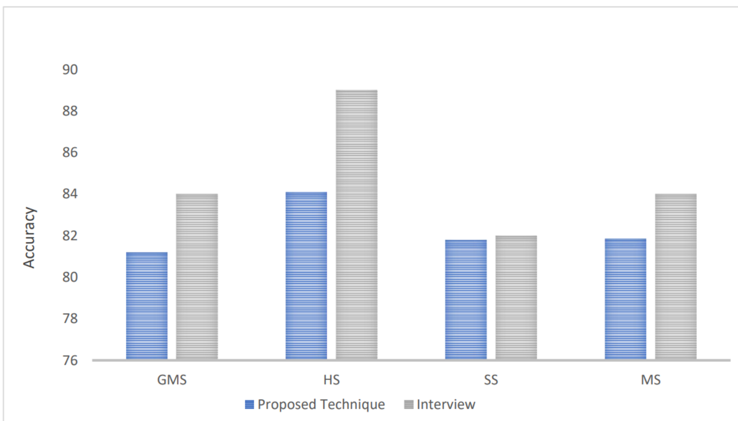

Figure 5: Comparison of proposed technique with interviews conducted

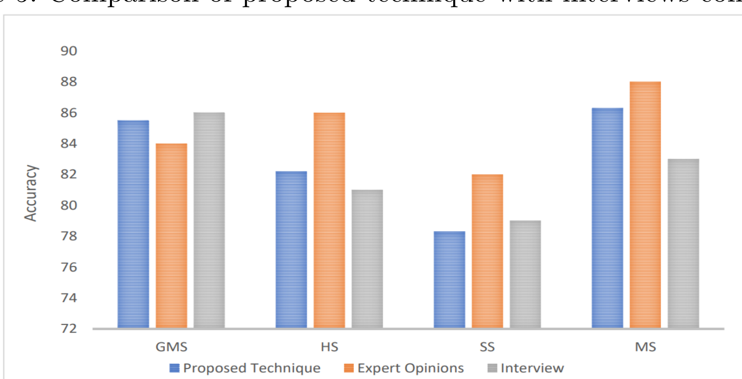

Figure 6: Comparison of proposed technique with the expert opinion and interview

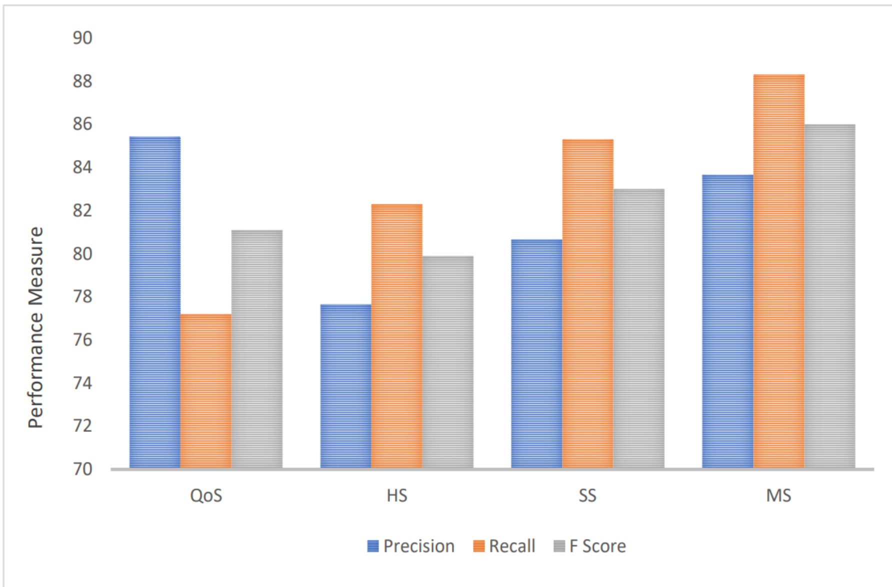

Figure 7: Experimental results in terms of precision, recall and f-score

5 Discussion

There was an agreement between our forecast and patient ratings regarding whether or not they would suggest a facility based on QoS. Our sentiment analysis prediction is accurate 78.5 percent to 87 percent of the time, depending on the classification method we utilized. GMS sentiment analysis is between 84 percent and 86 percent, HS prediction is between 82 percent and 84 percent, SS prediction is between 78 percent and 80 percent, and MS prediction is between 86 percent and 88 percent using these. Using ConceptNet, our suggested strategy yields promising findings and outperforms existing state-of-the-art methods such as surveys and interviews. Precision, recall, and F-score values are shown in Figure 7 for GMS, HS, SS, and MS.

6 Conclusion

This work predicts the sentiment analysis of patients' reviews related to their experience of medicinal services in some hospitals. This novel methodology is related to patient experience estimated by old methodologies i.e. surveys. This work contributes to understanding the patient reviews related to hospital QoS by analyzing their reviews posted on social media websites. Although there are some future possibilities by processing these reviews in real-time to get useful results. We improve it by refining the data polarity. At the same time, it is helpful to think about the general basic systems that are used in this process with different tools and platforms utilized for opinion mining and sentiment analysis i.e., SentiWordNet and WordNet Affect.

Future efforts to improve the characteristic of this algorithm include the improvement of its preprocessing ability to precisely predict the nature of comments given by the patients. Moreover, there is need to enhance the execution of sentiment analysis techniques, enhancement in the procedure to extract different types of free-content data on the Internet and analyze the connections between comments and clinical QoS. For example, there are a different features that are also added to improve the process by including a higher number of n-grams.

References

[1] F. Alemi, M. Torii, L. Clementz, D. C. Aron, Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments, *Quality Management in Healthcare* **21** (2012) 9–19. ⇒332

[2] M. R. Chauhan, A. Sharma, G. Sikka, The emergence of social media data and sentiment analysis in election prediction, *Journal of Ambient Intelligence and Humanized Computing* **12** (2021) 2601–2627. ⇒331, 338

[3] A. Chiche, B. Yitagesu, Part of speech tagging: a systematic review of deep learning and machine learning approaches, *Journal of Big Data* **9** (2022) 1–25. ⇒334

[4] J. S. Erickson, A. Viswanathan, J. Shinavier, Y. Shi, J. A. Hendler, Open government data: A data analytics approach, *IEEE Intelligent Systems* **28** (2013) 19–23. ⇒333

[5] G. G. Gao, J. S. McCullough, R. Agarwal, A, K. Jha, A changing landscape of physician quality reporting: analysis of patientsâ€™ online ratings of their physicians over a 5-year period, *Journal of medical Internet research* **14** (2012) 38. ⇒331

[6] M. Godovykh, A. Pizam, Measuring Patient Experience in Healthcare, *International Journal of Hospitality Management* **112** (2023) 103405. ⇒331

[7] F. Greaves, C. Millett, Consistently increasing numbers of online ratings of healthcare in England, *Journal of Medical Internet Research* **14** (2013) 94. ⇒332

[8] L. He, T. Yin, K. Zheng, They May Not Work! An evaluation of eleven sentiment analysis tools on seven social media datasets, *Journal of Biomedical Informatics* **1**32 (2022) 104142. ⇒335

[9] K. Herng Leong, D. Putri Dahnil, Classification of Healthcare Service Reviews with Sentiment Analysis to Refine User Satisfaction, *International Journal of Electrical and Computer Engineering Systems* **13** (2022) 323–330. ⇒331

[10] JA. M. Hopper, M. Uriyo, Using Sentiment Analysis to Review Patient Satisfaction Data Located on the nternet, *Journal of health organization and management* **29** (2015) 221–233. ⇒331

[11] J. W. Huppertz, P. Otto, Predicting HCAHPS Scores from Hospitals' Social Media Pages: A Sentiment Analysis, *Health Care Management Review* **43** (2018) 359–367. ⇒331

[12] M. R. Kanfoud, A. Bouramoul, SentiCode: A New Paradigm for One-time Training and Global Prediction in Multilingual Sentiment Analysis, *Journal of Intelligent Information Systems* **59** (2022) 501–522. ⇒331

[13] T. Lagu, S. L. Goff , N. S. Hannon, A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States, *The Joint Commission Journal on Quality and Patient Safety* **39** (2013) 1–7. ⇒332

[14] S. T. Lai, R. Mafas, Tentiment Analysis in Healthcare: Motives, Challenges and Opportunities pertaining to Machine Learning, *In 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2022 1–4. ⇒332

[15] B. Liang, H. Su, L. Gui, E. Cambria, R. Xu, Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks, *Knowledge-Based Systems* **235** (2022) 107643. ⇒335

[16] A. Nawaz, *Polarity Estimation and Aggregation in Aspect BasedSentiment Analysis,* PhD Thesis, International Islamic University, 2019 ⇒335

[17] A. Nawaz, T. Ali, Y. Hafeez, S. U. Rehman, M. R. Rashid, Mining public opinion: a sentiment based forecasting for democratic elections of Pakistan, *Spatial Information Research* **30** (2022) 169–181. ⇒334

[18] S. P. V. Rajeswaran, Bayesian analysis of ConceptNet relations on PubMed dataset, *Journal of Algebraic Statistics* 13 (2022) 2016–2025. ⇒335

[19] M. R. R. Rana, A. Nawaz, J. Iqbal, A Survey on Sentiment Classification Algorithms, Challenges and Applications, *Acta Universitatis Sapientiae, Informatica* **10** (2018) 58–72. ⇒331

[20] M. R. R. Rana, S. U. Rehman, A. Nawaz, T. Ali, A. Imran, A. Alzahrani, A. Almuhaimeed, Aspect-based Sentiment Analysis for Social Multimedia: A Hybrid Computational Framework, *Computer Systems Science and Engineering* **46** (2023) 2415–2428. ⇒335

[21] A. Rusiecki, Trimmed categorical cross-entropy for deep learning with label noise, *Electronics Letters* 55 (2019) 319–320. ⇒336

[22] H. Won, J. Han, M. S. Gil, Y. S. Moon, SODAS: Smart Open Data as a Service for Improving Interconnectivity and Data Usability, *Electronics* **12** (2023) 1237–2023. ⇒ 333
[23] H. Wu, C. Huang, S. Deng, Improving aspect-based sentiment analysis with Knowledge-aware Dependency Graph Network, *Information Fusion* **92** (2023) 289–299. ⇒ 331
[24] X. Xie, S. Ge, F. Hu, M. Xie, N. Jiang, An improved algorithm for sentiment analysis based on maximum entropy, *Soft Computing* **2**3 (2019) 599–611. ⇒ 336, 338

Received: July 31, 2023 • Revised: November 4, 2023