

Sentiment Analysis of Healthcare Quality

D. S. Panchal, S. S. Kawathekar, S. N. Deshmukh



Abstract: *In this work, concentrate on Hospital Service quality provided to patients through hospitals and hospital staff. Services are similar as the Doctor's reaction to the patients, instrument feature in hospital, area of expertise of Doctors, Compounder and Nurse. The patient expects good services from employees of hospital similar as from Doctor, ward boys, nurses, etc. These employees must take care of the patient's health, give needful medicine to the patient; the environment of the Hospital must be clean and comfortable for a patient. This paper concentrates on the patient's feedback about a hospital. We have taken patient's feedbacks from online website, social websites like Twitter, blogs, and Government healthcare-related website. These feedbacks are the form of comments available on websites. Then we need to analyze the comments, classify these comments as positive, negative and neutral by making use of sentiment analysis technique. Finally, based on results, we can decide which facility is in lower in quality, and recognize hospital which is better for patient's health improvement.*

Keywords: *Sentiment, Healthcare, patients, Twitter, Social Media, Hospital, Comments, feedback and Text.*

I. INTRODUCTION

There are many healthcare centers available in major cities, small towns, villages, and small societies. Patients admitted in the hospital want quality with careful service from the hospitals e.g. use of standard quality equipment in the hospital, service of employees, Doctors, ward boys, nurse, caretakers, cleanness and hospital charges. These service providers are important to analyze the quality of a hospital.

There are different complaints of patients about their health. Complaints are like body pain, patient's health issues, etc. Hospital service quality can be analyzed by the hospital owner using a sentiment analysis technique through commenting text. As per patients' suggestions they can improve their hospital service quality.

Collect online feedback commented text of patients from social websites like Twitter, blog, Healthcare website, government health websites, Reviews of healthcare, etc. These Websites are sources to collect online feedback data from various regions about hospital service quality. As we know that most famous social websites are Twitter, Facebook, and blogs in recent time, these websites playing interface between comment writer and reader. By using all websites, many users express their thoughts and opinion about a healthcare service, which he/she likes or dislikes.

Other users also take the benefit from this research paper.

Revised Manuscript Received on January 30, 2020.

* Correspondence Author

D. S. Panchal*, Department of C. S. and I. T., Dr. B. A. M. U. Aurangabad, India.

S. S. Kawathekar, Department of C. S. and I. T., Dr. B. A. M. U. Aurangabad, India.

S. N. Deshmukh, Department of C. S. and I. T., Dr. B. A. M. U. Aurangabad, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

This research paper gives information about patient feedback and pros and cons about health care. Some sample comments given by patients as a feedback on social websites are as follows.

Table- I: Sample Patients Comments

Sr. No.	Patient Comments
1.	"The Doctor in xyz hospital is very good",
2.	"Quality Staff, Efficient Service"
3.	"Gentleman intelligent efficient expert"
4.	"Great doctor very diligent and knowledgeable", down to earth despite his accomplishments!"
5.	"Best hospital and best path lab in Marathwada region. Less expensive treatments"
6.	"Outstanding result in chronic disorders, especially child and female disorders",
7.	"Excellent and smart doctor. Versatile and perfect solution to gynecological problems"
8.	"I'm thankful for a hospital, their staff, my Doctors for carefully handled me as a person instead of a patient".

The above comments are providing useful important information in short sentences. These comments can collect online as well as offline in the form of text.

Initially, the data are in the unstructured format. Therefore, our primary need is to analyze the data. There are some steps need to follow. The steps are as follows,

Pre-processing data, i.e. sorting, reducing and extract the desired result. On a government healthcare website the feedback of the patient is available in a question-answer format like yes or no. The data are available on some websites in 'csv' file format.

To analyze the data, a technique is used known as sentiment analyses. Using the Sentiment analysis technique, can classify text as positive, negative, neutral. For that, clean the text file; remove unwanted text, symbol or contents so that can analyze the comments without problems given by patients. After that, it can give a quality result which is needed to improve quality service to the patients [1].

A. Sentiment Analysis

Sentiment Analysis is linked to the people's opinions, thoughts, and belief. Natural Language Processing is used to implement and recognize the thoughts of the spokesperson or writer related to a particular subject matter. This is aim of this paper.

Hare, from unstructured raw data, can take out information; Because of this, it can be applied in the following regions.

- Surveys: analysis of authentic feedback of patients.
- Business and Governments: provide accurate and reliable information that can be used to make decisions and handle the sources increasing of negative response.
- Consumer feedback: analyze feedback reviews to improve patients experience and fulfillment.

- Health: For treatment of diseases and increase the drug searching process management of biomedical text is essential, to improve the healthcare quality [2].

B. Sentiment Analysis in Healthcare

To provide caring service to the needful patients must need to analyze patient's responses in text format. It is the first choice to improve the healthcare quality. Collect Surveys as question-answer like yes or no and inadequate patient's feedback report in a limited structure. For example, inadequate question-answer formats to collect patient feedback. Traditionally it is used, but these processes are very time consuming and very costly. It needs to select new tools and techniques to carry out new operations so that can get a much more preferred outcome.

Internet users are growing day by day in the current year. Because the internet is a way through which peoples are passing their post, thoughts into the all countries. These online facilities are also having some restrictions. Some social websites have some text restrictions. Initially, the text data is in an unstructured format. There are no rules to post online feedback to websites. Generally, these texts are open text information.

Therefore, after data collection, required to perform some operations on the data, e.g. preprocessing the data, sorting, merging, remove of unwanted text, symbols, etc. It needs to give a proper arrangement of the data and analyze the data as per requirement.

The United States found that 85% of senior citizens use the internet, 25% read others incident about the heat on an internet website or social blogs, and 11% people mentioned to online reviews of hospitals or hospital related websites.

Sentiment Analysis and Opinion mining with natural language processing, i.e. NLP has been difficult to analyze the consumer quality and response for health care. Sentiment analysis facilitates the content of natural languages, i.e. the words we use to be tested for positive and negative opinions or emotions. If linked to health care, these analytical methods can authorize analysis of textual information about a patient experience on a huge amount. Alemiet. al. Have suggested to use of sentiment analysis comments as live patient surveys.

After that, the work assures by Alemi and his group, by analyzing a large content of open-text comments on the NHS website. Most of the information about a patient's treatment experiences from NHS Website England. These comments are millions of visits in a day, from 2 years data, each hospital common mean is 69 reviews. These reviews include common incident and ratings of the Likert - scale feature of health care. This demonstrates an option for probable trial to review the accuracy of sentiment analysis techniques proceed to the patients' self-qualitative ratings.

Need to analyze whether there is a link with usual actions of the patient's experience, if sentiment analysis technique is considered as an applicable tool for evaluating the quality of health care. For that reason, we need to compare sentiment analysis result to a national survey of patients, at the hospital level.

Current time, several hospitals or doctors' offices use sentiment analysis to find out the patient satisfaction level. The internet user increased daily and uses the internet to give their feedback about a particular doctors or hospitals. This

paper could be used to analyze the healthcare situation through sentiment analysis [3].

II. STATEMENT OF THE PROBLEM

Improvement of hospital quality is needful. Some hospitals are very low in quality services. No cleanliness, careless staff, many doctors are not able to give proper treatment to serious patients. Some hospitals are very costly, common people are not able to pay the expensive fees.

In many hospitals well-equipped infrastructure is not available, e.g. shortage of standardized equipment, shortage in medicine, etc.

Because of all these things, most patients are unsatisfied with healthcare quality services. Sometimes patients are not able to reach the desired hospital in their needful critical situation. Hospitals are not able to give proper service to the needful patients.

To solve these issues must need to analyze the health care services related issues. This paper shows the hospital service quality issues based on the patient's feedback and his experience at the healthcare center [4].

III. RELATED WORK

Take out the sentiment from healthcare related data; can achieve the best quality result for improvement in a related sector as healthcare, including views from beneficiaries such as patients. However, the essential complexity of biomedical tests (e.g. Due to the use of abbreviations and frequent misspellings) make the health information services one of the most difficult fields for text analytics to be practical. As a result, there is a large volume of research intended to provide enhancement in this area, somewhat due to the benefits related to improved health care for human being e.g., focus on calculating heart failure by creating an NLP, to recognize the signs and indication related to the status. To find early heart failure signs. They look problems while transferring data from an Electronic Health Record because of spelling errors, a common problem that creates confusion in the healthcare domain. Since part of this research is to ensure healthcare-related reactions are analyzed appropriately, it is believed sentiment analysis could help the patient-doctor communication and improve the patient's experience. It could also help with the analysis of EHR data. A basic thought about sentiment analysis is subjectivity from subjective texts is expected to openly convey feelings and beliefs that build an opinion. Hence, various studies focus on understanding and recognizing subjective sentences. Therefore of the apparent link among subjectivity and opinions, researchers are likely to fail to notice objectivity in the belief that there is no essential loss of information.

However, Benamar et al. provides evidence to justify that sentiment can exist in both types of sentences by examining different combinations of subjective and objective sentences. As explained in reference number [4], the structure recommends a feedback method wherein, sentiment analysis is carried out from surveys and tweets based on usual health issues among adult women in India.

The social opinion on current health problems is analyzed, and events are taken to create responsiveness using email, SMS, blog, forum posts or website posts.

Center of attention on an analysis of the opinions and subjects confer in the forum. The system's working is briefed with an example. Many adult women around thirty years in the Asian country are making certain with Thyroid problems in current years, which are mostly put in by stress factor. Sentiment analysis is carried out during this variety and if positive emotions are declared, then awareness programs can be initiated for Thyroid issues and Stress Control. Periodically current problems are initiated and Sentiment analysis is performed, consequently awareness initiatives are created.

It helps improving the performance of healthcare. This alertness initiative helps in concentrating on current health issues that area unit widespread and if the reach of such alertness programs is best, then the alertness created may have a high impact on older women. Sentiment analysis in health care makes use of natural language software to classify and evaluate written and spoken comments by patients about their healthcare experience and ideas.

Generally, we can say that sentiment analysis objective is to predict the focus of a speaker or a writer with respect to the related aspect or the overall contextual polarity of a document, e.g. his or her analysis or evaluation, affective state (i.e., the emotional state of the author when writing), or the intended emotional communication (that is to say, the emotional effect of the author wishes to have on the reader). The initial and important piece of work using sentiment analysis is categorizing the data of a given text in the document, sentence, feature, and aspect level.

A statement or units feature, or aspect is positive, negative or neutral and emotional states like "angry," "sad," and "happy".

As mentioned in [4] 4, 04,065 tweets intended for to 2349 hospitals in the US for one year time were classified along patient knowledge with the help of machine learning techniques. Through Sentiment analysis these tweets were analyzed. [5] In this reference paper 11,602 tweets classified related to patients experience topics.

Finally, hospitals having more than 50 or 50 patient experience tweets were surveyed to analyze how they use Twitter to interact with patients. Key results roughly half the hospitals within the United States have a presence on Twitter. The tweets are showing toward these hospitals, 34,725 i.e. 9.4% were related to patient experience and covered diverse topics.

Analysis of tweets restricted to hospitals with ≥ 50 patient discovered that they were additional active on Twitter, additional probable to be below the national median of health care patients ($p < 0.001$) and above the national median for nurse/patient ratio ($p = 0.006$), and to be a nonprofit hospital ($p < 0.001$). After adjusting for hospital characteristics, the author observed that Twitter sentiment was not linked to the Hospital Consumer Assessment of Healthcare Providers and Systems ratings, although there was a less relation with 30-day hospital readmission rates i.e. $p = 0.003$. [4] Shows statistical analysis on Pearson's correlation to assess the linear relationship between numeric variables, Fisher's exact test to compare proportions between categorical variables,

and a two-tailed independent t-test to compare the means of Quintiles. Bonferroni correction was used to adjust for multiple comparisons.

Multi variable linear regression was used to adjust for potential confounders such as region, size, bed count, profit status, rural/urban status, teaching status, nurse-to-patient ratio, a percentage of patients on Medicare and percentage of patients on Medicaid. Twitter account holders were calculated in August 2014. The additional Twitter analysis was that the total number of patient experience tweets was received during the analysis time. Whether the hospital is a single user of, which is against the distribution of his health care information for keeping confidentiality, for that a Wald test was used to test for tendency importance.

Last ten years, patient experiences found increasing curiosity, highlighting the importance of integrating patients' need and viewpoint into care delivery. As healthcare concentrating on patients, healthcare partners need to be able to evaluate the report and improve outcomes that are meaningful to patients. These outcomes can only provide by patients, and therefore the system is needed to capture outcomes reported by patients and make possible the use of this information at both an individual patient level and the population level. Usual patient experience surveys i.e. the Hospital Consumer Assessment of Healthcare provider and Systems (HCAHPS) are an ordinary process designed to assess the patients' view of the quality of their own health care.

The main disadvantage with these surveys is an important time consuming several months before authorized data are free, making it difficult for patients and other concerned partners to be informed about current opinions on the quality of a given organization. Furthermore, these surveys usually have low response rates, increase concerns about likely response and selection partiality in the results [6].

A. The Value of Sentiment Analysis

The value of sentiment analysis found in its capability to transfer unstructured text into structured data that can be quantified and analyzed, permit a hospital or other provider to grow above the amazing details provided individual comments to a deeper analysis of experience and emotions – the drivers of survey scores. Specifically, sentiment analysis can help to develop an actionable idea, i.e. early warnings or alerts, root-cause analyses, trending over time, and identifying the intensity of reactions and feelings to improve focus on the most important issues.

Measure the impact of targeted interventions over time. Once a key issue has been identified, using sentiment analysis and intervention designed and implemented, the results of the action can be measured by trending the volume of comments and their sentiment scores over subsequent periods [7].

IV. METHODOLOGY

The possible research methodology will be first to manually comment on Patients experience. Comments provided by users as negative or positive.

Then it will use two ways one is by using key-word based approach and second is Machine based approach [8].

A. Keyword-based approach

Can create two groups of keywords one class of positive keywords and another class of negative keywords. Then, based on these keywords, will classify patients' experience comments are as positive, negative or Neutral.

The keyword-based approach is defined as the method that repeatedly recognizes a set of terms that best describe the subject of a document.

Various terminologies are used in analyzing the terms that show the most related information included in the document like Key phrases, key segments, key terms or keywords.

Extracting a small group of units made of more than one or single terms, from a single document is a significant difficulty in Text Mining, information retrieval, and natural language processing. The keyword is commonly used to permit queries with information retrieval systems as they are simple to define, revise, remember and share. When comparing mathematical signatures they are independent of any corpus and can be used entirely different databases and information retrieval systems.

Keywords have also used to get better the functionality of IR systems. In another definition, these related extracted keywords can be applied to create an automatic index for a document collection or on the other side can be used for document demonstration in classification tasks.

The capturing summary of the document is the main assignment of many information retrievals. NLP applications include automatic indexing, automatic summarization, document management, high-level semantic description, text, website or document classification or clustering, vice versa classification retrieval, building specific dictionaries, name entity recognition, topic detection, tracking, etc. [9].

B. The machine learning approach

Need to develop a classifier that will classify patient experience comments as positive or negative. The classifier will be trained because of some manually classified positive and negative comments. Therefore, it will be used patient experience of comments received from Twitter, health blogs, etc.

Training set and test set machine learning techniques used. The training set has input characteristic vectors and their equivalent class labels. Classification model created using these techniques, which tries to classify the input characteristic vectors into equivalent class labels. Next, a test set is used to authenticate the form by predicting the class labels of unnoticed characteristic vectors.

Different machine learning techniques is utilized e.g. Naive Bayes, Maximum Entropy, and Support Vector Machines classify reviews or text. In sentiment classification, there are some other features used for classification similar to a presence of terms, a frequency of terms, negation, n-grams, and parts of speech. These features can be used to find out the linguistic orientation of words, phrases, sentences, and documents. Polarity is either positive or negative also known as linguistic orientation.

Domingo's et al. From different classifier Naive Bayes classifier works fine. This is unexpected as the basic hypothesis of Naive Bayes is that the characteristics are

independent. Zhen Niu et al. Characteristics selection, weight calculation, and classification are introduced as a new form in which competent ways are used for the new model is based on a Bayesian algorithm. Here, load of the classifier is used to for representative attributes and distinctive feature. 'Representative attributes' is the feature that displays a class and 'distinctive characteristics' are the information that helps in distinctive classes. The probability calculated using those weights of each classification and will get better Bayesian algorithm. Barbosa et al. Classifying tweets it is intended a 2-step mechanical sentiment analysis method. In developing classifiers they make use of a noisy build upset to decrease the labeling trial. First step they categorize tweets into subjective and objective technique. Then, tweets for subjective are categorized as positive and negative tweets.

Celikyilmaz et al. Has developed the clustering method based on normalized noisy tweets pronunciation. Common tokens assigned to clustered text for utterance based word cluster, words having the same utterance. Some text processing method used like to convey related token for numbers, HTML links, user identifiers and objective organization names for normalization. After normalization, they used probabilistic models to recognize polarity lexicons.

With these polarity lexicons as features and obtained a less error rate while performed classification using BoosTexter classifier. Wu et al. Presented a useful probability model for Twitter sentiment analysis. If @tweetusername is initiated in the body of a tweet, it is valuable action and it put into useful probability. Any tweet that starts with @tweetusername is a retweet that produces and motivated action and it devote to a motivated probability. They found that there is a strong connection between these probabilities. Pak et al. Created a twitter corpus by mechanically collecting tweets using Twitter API and mechanically annotating those using emoticons. A Machine learning technique in which an algorithm learns to classify comments into dissimilar class from a set of data by using Weka data mining software. Most of the previous research work used the Weka software and it gives a precise classification result, along with health care.

For confirmation of correctness, compare outcomes to quantitative ratings collected from the same like individual patients on a Likert scale. Then analyze through question-answer reaction, i.e. "What I liked", "what could have been improved", and "any other comments", after that the calculation will obtain from the patient about the advices of hospital or not, i.e. whether the hospital is clean or dirty, and whether they were treated with dignity and respect [10].

In reference to number [2] the algorithm was initiated by affecting the comments and ratings regarding hospitals left on the NHS alternative website from year 2008, 2009, and 2011. Total 13,802 as a learning set. 2010 Data be utilized to check the calculate correctness of the procedure. Comments near about 6412 patient data were accessible for that year.

V. SOURCES OF DATA

There are different sources of data, collect patients' comment from social websites, healthcare websites, blogs, etc.

For this work, collected a database from a website known as <https://www.ratemds.com>. This website contains detailed information about Doctors followed by their specialty and experienced patient comments. These comments contain feedback information about the Doctor and hospital. There are many physicians detail are added to this website, including all over the world's physicians.

“ratemds” website also included Indian Physicians of different Indian cities with their specialty. To find a particular physician into “ratemds” website you must need to select a doctor’s specialty, the city followed by Male or Female. After that, here we can get physicians profile as well as patients feedback.

From this website, collected more than sixty physician’s feedback text commented by patients. This text data are collected as per the specialty of physicians. With this data we have stored in an excel file and convert to ‘CSV’ format for further processing [12].

VI. PRE-PROCESSING

Data pre-processing is a process in which processing is performed on raw data to organize it to remove noise or unwanted text for the purpose of data cleaning and for better classification. Pre-processing is like to removing unwanted text, punctuation, NA values, numbers, etc. Fig. 1 shows the details of pre-processing steps one by one.

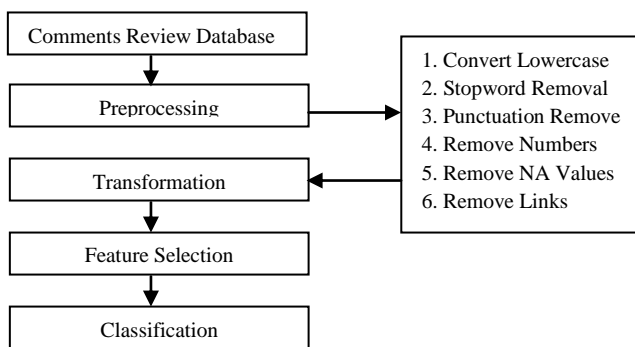


Fig. 1. Pre-processing

A. Lowercase

Convert all capital letters to lowercase.

E.g. A-a.

B. Remove Punctuation

The Process of removing Punctuation marks from a text.

E.g. comma, Full stops, exclamatory sentences, etc.

C. Remove Stop word

The Process of removing Stop words from a text.

E.g. a, and, but, how, what.

D. Remove NA Values

The Process of removing NA values from a text.

E.g. NA

E. Remove Numbers

The Process of removing numbers from a text.

E.g. 1, 2, 3, 0, -1.

F. Remove Links

The Process of removing links from a text.

E.g. www.abc.com [13]

VII. RESULTS AND DISCUSSIONS

First of all, need to collect patients commented text from a website known as “ratemds.com”. As mentioned in data sources this text comments selected as per specialty of Doctor, location, and gender. Then store this text into excel file. After that, convert this excel file into a CSV file format. And after that, perform the different pre-processing operation on this ‘c.s.v.’ file.

As mentioned in the pre-processing section Convert all text into lowercase, remove all punctuation marks from the text, remove stop words, remove NA values, remove numbers, and remove Links, etc.

As Discussed below the detailed architecture to predict the result of the Healthcare on the basis of patient’s feedback, i.e. commented online text [14].

A. System Architecture

Following Figure 2 Shows the system architecture of a detail flows diagram.

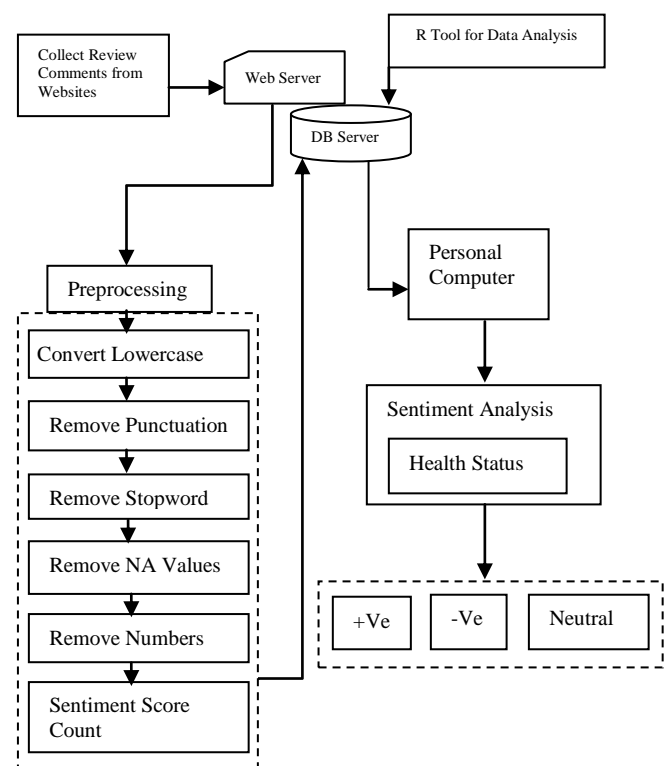


Fig.2. System Architecture

This system architecture shows the flow of pre-processing along with sentiment analysis and R tool for data analysis [15].

- 1) *R-Tool*: It is a tool used for R-Programming. Using this tool we can develop a program to predict the sentiment score of text.
- 2) *Sentiment Analysis*: Using this technique we can find the sentiment score about a patients comment using the R - program
- 3) *Positive text*: this contains a positive word database having all positive words with synonyms.
- 4) *Negative Text*: in this text file negative word is included along with synonyms.

B. Sentiment Score

After pre-processing need to count the sentiment score. To count sentiment score R – tool is used. Then create a program in R-language, with the help of this R-program, count the sentiment score of each comment comparing with a Positive and negative text file. This positive and negative text file collected from Github website “Twitter sentiment analysis tutorial”

After comparing the positive and negative text file with patients commented database about the physician, then get sentiment score. This score is counting the positive and negative words from a sentence and after that, it will show the sentiment score [16].

In this table II decided that those comments having scored more than 0 that comment is considered as a positive comment, comments less than zero considered as negative and those comments having scored exactly zero that comments is considered as neutral comment.

Table-II:Sample Sentiment Score of patients’ comments

Sr. No.	Score	Comments
1	0	ONE OF THE UNIQUE HOSPITAL IN THE REGION PROVIDING ALL RESPIRATORY CARE UNDER ONE ROOF.
3	-1	Confident, Competent, intelligent and 'sensitive' Anesthesiologist of poor and risky patients.
6	3	Very polite, gives time to patient, explains precisely.

Above Table II Shows the sample sentiments score of a comment which is predicted using the R programming in R-Tool. In this table, see that score 0 means comments are neutral, score -1 means negative comment and score 3 means positive comment.

Table- III: Sentiment Score of comments

Sr. No.	Number of Comments	Sentiment Score
D1	24	(7, 1, -2, -3, 1, -5, 6, 4, 1, -3, 6, -2, -2, 3, 2, 3, 12, 0, 1, -1, 2, 1, -1, 0)
D2	24	(3, 2, 5, 4, 5, 5, 1, 1, 1, 3, -10, 6, 0, 3, 2, -1, 2, 3, 2, 1, 1, 3, 4, 1)
D3	23	(0, -2, -23, 1, 1, 2, -1, 2, 8, 4, 1, 1, 0, 2, 3, -1, -1, 0, 3, -2, 1, 2, 6)
D4	18	(12, 4, 1, 2, 2, 2, -7, 1, 4, 2, 4, 2, -1, 6, 0, 2, 1, 0)
D5	14	(3, 2, 2, 1, 0, 1, 2, 2, 1, 0, 2, 0, 2, 1)
D6	12	(4, 3, 1, 6, 2, 2, 7, 3, 3, 5, 5, 0)
D7	12	(21, 2, 6, 2, -2, 18, 10, 2, 12, -1, -1, 5)
D8	10	(5, 9, 2, 5, 1, -2, 4, 0, 1, 5)

Above table III shows the sentiment score of each commented feedback of a patient about a physician, and each comments sentiment score is separated by comma. Total feedback, comments given by patients of Doctor D1 are 24. Each comments sentiment score is shown in sentiment score column of Table III. i.e. sentiment score of comment number one of Doctor D1 is ‘7’, sentiment scores of comment number two of Doctor D2 is ‘2’, sentiment scores of comment number three of Doctor D3 is ‘-23’ etc.

Total comments and extracted comments in classified manner are shown in Following Table IV. This table contains

doctor number, total comments, positive comment out of total comments, negative comments out of total comments and neutral comments out of the total comments. These comments are showing the opinion of patient about a physician and classified as positive, negative and neutral opinion about a physician.

Many positive, negative or neutral comments are given. Here in this table, we can observe that most of the comments are positive response about a physician, a negative response is very less as compared to positive response [17].

In order to maintain anonymity we use D1, D2, D3 etc. represent the doctors.

Table IV:Classified commented feedback of patients

Sr. NO.	Number of Comments	Negative Comments Count	Positive Comments Count	Neutral Comments Count
D1	24	8	14	2
D2	24	2	21	1
D3	23	6	14	3
D4	18	2	14	2
D5	14	0	11	3
D6	12	0	11	1
D7	12	3	9	0
D8	10	1	8	1

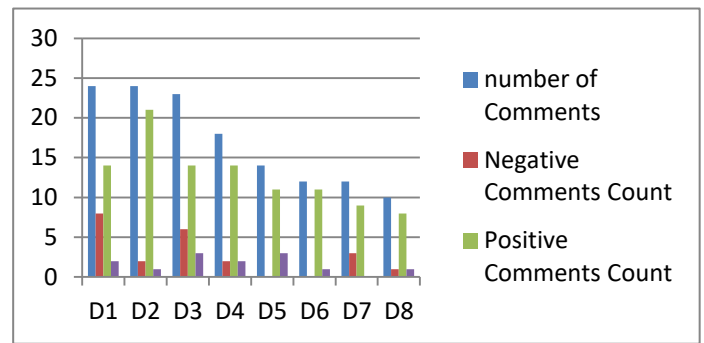


Fig.3. Graphical Representation of Table III

Figure 3 shows the graphical representation of Table IV. Graphically, it can be found that the positive response of patients about a physician is more than negative response as well as neutral response.

Table- V: Classified commented feedback of patients in Percentage

Sr. NO.	Percentage of Negative	Percentage of Positive	Percentage of Neutral
D1	33.3	58.33	8.33
D2	8.3	87.50	4.17
D3	26.1	60.87	13.04
D4	11.1	77.78	11.11
D5	0.0	78.57	21.43
D6	0.0	91.67	8.33
D7	25.0	75.00	0.00
D8	10.0	80.00	10.00

Table V shows the percentage of patient’s opinion on the basis of above table III. In table IV Percentage wise comments are shown about a physician.

As per table IV Percentage of Positive opinion is more than negative and neutral opinion of maximum patients about a physician. Some patient’s negative opinion found the null.

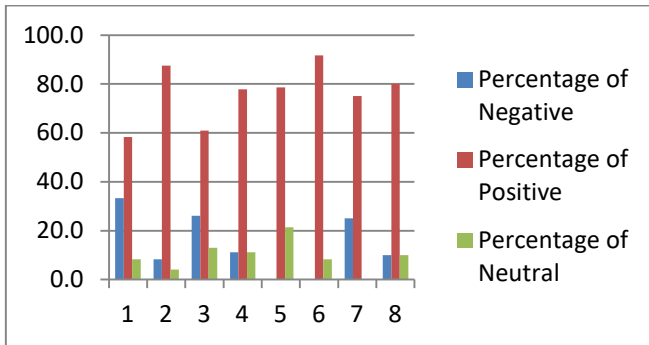


Fig. 4. Graphical Representation of Table IV

Above Figure 4 showing the graphical representation of table number IV. It is found that percentage of positive response on a patients graph is higher than negative and neutral graph. The percentage of the negative and neutral graph is very less as compared to positive graph [18].

VIII. RESEARCH LIMITATIONS

The analysis of this research paper has some limitations, because not all peoples are a continuous user of the internet. Cannot able to reach to those peoples who are not using the internet. Therefore, cannot receive feedback data of patients, those who are not aware of the internet. It may difficult to prove that doctors are assuming comments on websites are truthful.

Very less number of physicians are available on www.Ratemds.com website, because of this not able to collect the maximum physician’s feedback. There may be a negative effect on these results. Another is we cannot able to predict the how much truthful feedbacks in comments posted by patients. However, availability of patients’ comments is limited, for some doctors affecting the highest accuracy in prediction. [19].

IX. CONCLUSIONS

This research paper has shown Sentiment Analysis techniques successfully identifying free-text comments relating to healthcare improvement of experienced patients, within a large data set. The techniques used in this study facilitate qualitative analysis in resources of comments. This study provided visions into significant associations found between participants’ comments relating to the quality of their healthcare.

As a future scope, the patients comment from social media will be analyzed for predictions.

ACKNOWLEDGMENT

We would like to thank the Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India for supporting this work.

REFERENCES

- G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
- Georgiou Despo, Andrew MacFarlane, “Extracting Sentiment from Healthcare Survey Data: An evaluation of sentiment analysis tools”, Department of Computer Science City University London, U.K., Tony Russel-Rose, UXLabs, Guildford, U.K, July 28-30, 2015, pp. 1-10.

- Greaves Felix, Daniel Ramirez-Cano, Christopher Millett, AraDarzi, Liam Donaldson, “Use of Sentiment Analysis for Capturing Patient Experience From Free-Text Comments Posted Online”, *Public Health England*, Imperial College London, London School of Hygiene and Tropical Medical, pp.1-17.
- Hopper Anthony M., Maria Uriyo, “Using sentiment analysis to review patient satisfaction data located on the internet”, Department of Health Systems Administration, Georgetown University, Washington, District of Columbia, USA, 22 August 2013, pp. 1-15.
- Jeyabalan Jeyalakshmi, SindhuManivasagam, Poornimathi Krishnan, SreesubhaSoundararajan, AnithaJaikumar, "Sentiment analysis for improving the healthcare system for women", Department of Information Technology Rajalakshmi Engineering College Thandalam, Chennai, India.
- Hawkins Jared B, John S Brownstein, Gaurav Tuli, Tessa Runels, Katherine Broecker, Elaine O Nsoesie, David J McIver, Ronen Rozenblum, Adam Wright,7,8 Florence T Bourgeois, Felix Greaves, “Measuring patient-perceived quality Of care in US hospitals using Twitter”, *HealthMap BCH3393*, 300 Longwood Ave, Boston, MA 02115, USA; jared, October 2015, pp. 404-413.
- Deng Yihan, MatthaeusStoehr, Kerstin Denecke, “Retrieving Attitudes: Sentiment Analysis from Clinical Narratives”, University of LeipzigSemmelweisstr. 14 Leipzig, Germany. ENT Clinic University Hospital Leipzig Liebigstr. 10 Leipzig, Germany., ICCAS University of Leipzig Semmelweisstr. 14, Leipzig, Germany, 11th July 2014, pp. 12-15.
- Suchdev Riya, PallaviKotkar, Rahul Ravindran , Sridhar Swamy, “Twitter Sentiment Analysis using Machine Learning and Knowledge-based Approach”, *Computer Engineering VES Institute of Technology*, University of Mumbai, Mumbai, India, October 2014, pp. 36-40.
- By Richard B. Siegrist, CPA, “The Science of Emotion Sentiment Analysis Turns Patients Feelings into Actionable Data to Improve the quality of care”, Jr. MS, MBA, CPA, Chief Executive Officer; and Susan a Test Topic”, Department of Health Science, Maddan, MS, Vice President of Product Analytics, Press Ganey Associates, Jan/Feb 2011, pp. 27-35.
- Nishantha Medagoda, SubanaShanmuganathan, “Keywords Based Temporal Sentiment Analysis”, *School of Computer and Mathematical Sciences Auckland University of Technology Auckland, New Zealand*, pp.1-9.
- Ahmed Ali, WalidMangdy, and Stephan Vogal, “A Tool for Monitoring and Analyzing HealthCare Tweets”, *Qatar Computing Research Institute Qatar Foundation Doha, Qatar*, 1 August 2013, pp. 1-4.
- Paramjoth Kaur Chahal, “Sentiment Analysis of Data Collected from Social Media for Improving HealthCare”, *The School of Engineering & Computing Sciences Texas A&M University-Corpus Christi Corpus Christi, TX*, pp.1-38, 2015.
- Sujith Jayaprakash, Balamurugan E., “A Comprehensive Survey on Data Preprocessing Methods in Web Usage Mining”, *International Journal of Computer Science and Information Technologies*, Vol. 6 (3), pp.3170-3174, 2015.
- Paramjoth Kaur Chahal, “Sentiment Analysis of Data Collected from Social Media for Improving HealthCare” GRADUATE PROJECT PROPOSAL, The School of Engineering & Computing Sciences Texas A&M University-Corpus Christi Corpus Christi, TX, pp. 1-38, 2015.
- Hein Htet, Soe Soe Khaing, and Yi Yi Myint, "Tweets Sentiment Analysis for Healthcare on Big Data Processing and IoT Architecture Using Maximum Entropy Classifier", University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Myanmar, pp.28-38, 2019.
- Tanveer Ali, David Schramm, Marina Sokolova, Diana Inkpen. “Can I hear you? Sentiment Analysis on Medical Forums”, *International Joint Conference on Natural Language Processing*, Nagoya, Japan, pp. 667–673, 14-18 October 2013.
- Ms. GaurangiPatil, Ms. VarshaGalande, Mr. VedantKekan, Ms. KalpanaDange, “Sentiment Analysis Using Support Vector Machine”, *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, Issue 1, pp.2607-2612, 2014.
- Ms. GaurangiPatil, Ms. VarshaGalande, Mr. VedantKekan, Ms. KalpanaDange, “Sentiment Analysis Using Support Vector Machine”, *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, Issue 1, pp.2607-2612, 2014.
- Felix Greaves,1 Antony A Laverty,2 Daniel Ramirez Cano,1Karo Moilanen,3 Stephen Pulman,3 Ara Darzi,1 Christopher Millett,2, “Tweets about hospital quality: a mixed methods study”, *BMJ Quality & Safety Online First*, pp.1-9, 19 April 2014.



AUTHORS PROFILE



D. S. Panchal Research Student, M.Sc. (I.T.), M.Phil. Computer Science, Ph.D. Computer Science (Pursuing), at Department of Computer Science and I.T., Dr. Babasaheb Ambedkar Marathwada University, Aurangabad. Till now seven research papers published. The research area is in Sentiment Analysis, Text Mining, Data Mining, and Big Data Analysis.



S. S. Kawathekar Assistant Professor, M.Sc. Computer Science, Ph.D. Computer Science. Currently working as Assistant Professor at Department of Computer Science and I. T., Dr. Babasaheb Ambedkar Marathwada University, Aurangabad and having experience of around Twenty Five years in teaching for Graduate (B.Sc., B.C.S. And Post Graduate (M.Sc.)). Member of IETE, Research Area is Medical Image Processing, Computer Graphics, Data Mining, Multibiometrics. And Twenty Seven Research papers are published in National and International Journals and conferences.



S. N. Deshmukh Professor, M.E., (CSE) Ph.D. Currently works as Professor in Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad and having experience of around Twenty Two years in teaching for postgraduate (M. Tech, M.Sc. And MCA) and graduate courses (B.E., B. Tech.) and also have an experience of software development using .NET, Oracle, J2EE, XML, Java Scripts, PHP and MySQL. He has also received the appreciation from University for development of Administration software. He has also conducted more than 10 workshops on above technology. Research Area is Data Mining, Text Mining, Social Media Data Analytics, Sentiment Analysis, Opinion Mining, and Intension Mining. More than thirty four research Paper Published in National, International Journals and Conferences.