Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7778551

PATIENT HEALTH CARE OPINION SYSTEMS USING ENSEMBLE LEARNING

C.Arulananthan[1]*, K. Sujith[2], D.Vaishnavi[3]
[1*] Department of Computer Science, Annai College of Arts & Science, Kumbakonam, Tamilnadu, India. (Affiliated to Bharathidasan University, Tiruchirapalli) E-mail: arul.9943742532@gmail.com
[2] Department of Computer Science, Annai College of Arts & Science, Kumbakonam, Tamilnadu, India. (Affiliated to Bharathidasan University, Tiruchirapalli) E-mail: mailme.sujith@gmail.com
[3]Dept. of CSE, SRC, SASTRA Deemed to be University, Tamilnadu, India. E-mail: vaishume11@gmail.com
*Corresponding Author: C.Arulananthan
*Research Scholar, Dept. of CS, Annai College of Arts & Science, Kumbakonam, Tamilnadu, India. E-mail: arul.9943742532@gmail.com

Abstract
The patients' experience is considered a dominant reputation in the hospital administration and medical fields. Online patient reviews are recognized as an important criterion for evaluating hospital service quality and performance. The classical approach of evaluating service excellence is often found to be tedious. But with machine learning classifiers and opinion mining techniques the data assessing, and evaluation is made casual and its saves time. Currently, patient satisfaction and quality of service for patients in hospitals plays a major role in health care sector. In this paper a novel Ensemble Model is proposed to Analyze Patient Health Care Opinion Systems. The Classification models are used to classify patients' feelings as positive, negative, or neutral using a machine learning approach to predict superlative models in data analysis. Ensemble techniques are used to analyze the opinions classified by the model, and the recommendation for health care is analyzed based on sentiment polarity.
Keywords: Healthcare system, Ensemble learning, Data Analysis

1. INTRODUCTION

In recent times, online reviews help patients make decisions by observing the tweets based on hospital services, medicinal care, hospital atmosphere etc. However, the efficiency primarily lies on how effectively these data going to be utilized in favor of patient decision- making support system as it has associated with sentiments, feelings [1]. The key point is to capture patients' needs and comfort level during the stay, yet reviews collected during the visit exhibited unseemly contents because of irrelevant questions posted. Instead, the online mode review collection form has been introduced to overcome trouble. Nonetheless, there is no significant success rate achieved; either patient turned up with poor responses or were ignored altogether. Nowadays, the preferred platform to express the opinion is through social media, where the patients can respond instinctively in the form of assessments, tweets, ranking, etc, The information extracted from these tweets facilitated us to analyze the data, stem the hidden patterns and build an effective recommendation system for the health care sector [2].

Human beings' welfare is well-thought-out to be personals as it changes based on individuals need and livelihood as it is an immense area of study in sentiment analysis [3]. Nowadays, it has become a challenge to handle the society's anxiety on negative sentiment pertaining to diseases, treatments, injuries etc, [4]. At the same time, it is also a fact that social media forums and discussions derive many optimistic inferences.When these views are analyzed systematically, it creates an opportunity to act upon customer reviews by increasing goodwill [5]. Using machine learning techniques, the accuracy prediction in hidden patterns is simplified with different classification approaches [6]. This work will examine the online forum ratings of various hospitals in and around chennai, India. Classification models are used to classify patient's feelings as positive, negative or neutral using machine learning approach for predicting superlative model in data analysis.

The very reason behind selecting this topic is to provide an effective information system to the health care industry based on the tweets posted. In the pre-processed reviews, various Machine Learning classification models such as Logistic Regression [7], Support Vector Machine [8], Random Forest Classifier [9], and K-Nearest Neighbor [10] were used. A classification report is generated using the test models. The scores generated were precision, recall, F1- measure and accuracy. The accuracy has been verified using Cross Validation method, it reveals that Random Classifiers as the best accuracy model over other models studied after the ensemble technique. This paper is further consists of: section 2 discusses the proposed methodology, the data collection description, and steps in designing these models using feature extraction and sentiment classifier approaches. Sections 3 discusses the proposed methodology's result analysis along with its validating methods and finally, section 3 infers conclusion on materials and techniques, results and discussions.

2. PROPOSED METHODOLOGY

Patient satisfaction is a crucial factor influencing the global healthcare industry. According to one study, a US-based hospital chain experienced a problem with patient satisfaction, which impacted the hospital's growth [11]. To address this issue, machine learning techniques are used to analyze patient opinions about the services provided. These machine learning analyses helped to regain their business by identifying the root cause problem and providing timely support in resolving the errors, ensuring patient satisfaction. These days, it has become a challenge due to increase in healthcare chains and operations involved in running these healthcare sectors, where patients demand more facilities in terms of quality services. It makes the higher authority focus on these factors and correct the root cause problems to fetch patients' satisfaction. This research work aims at proposing some novel approaches based on below contributions like:

1. A general approach towards sentiment analysis for any healthcare reviews.
2. Collection of nearly 5000 reviews through online Google reviews using outscraper.com for six hospitals in and around Chennai.
3. The recommendation for the hospital is based on the relationship among the reviews.
4. Applying Sentiment classifier approaches using ensemble technique for multiclass model building and performance metrics.
5. Categorization of the reviews as positive, negative and neutral.

Patients can compare the hospital's merit based on the recommendation percentage obtained using the sentiment polarity of the hospital reviews. The sentiment analysis problem is framed to provide a full understanding of user opinion based on various characteristics and analyze the opinion using NLP. The techniques used in healthcare for classifying the text are positive, negative, and neutral. It is done using trait-based sentiment analysis methods. This proposed research work targets in evaluating the dependencies between user ratings and textual representations. The model consists of three main steps namely Data pre-processing, sentiment extraction and bag-of-words representation and sentiment classification via classifier modelling it.

2.1 DATA COLLECTION AND PRE-PROCESSING

Firstly, the data is extracted from the web forum of different healthcare centers in chennai, India. Data was obtained using Outscraper.com's API created using a Python script. Initially, ratings were collected in Excel format, which was later converted to CSV format. While original and repeat reviews are included, independent reviews are excluded and manually classified each review as positive, negative, or neutral, including original tweets and retweets.

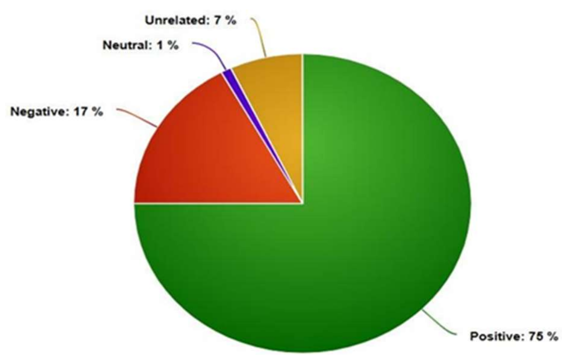

Fig. 1 Distribution of dataset based on class labels

2.2 PROPOSED ARCHITECTURE

After excluding unrelated reviews, original and repeated reviews are included and manually categorized each review as positive, negative, or neutral. The feedback obtained from six hospitals in the Chennai locality was considered. As previously stated, a total of 5000 reviews were collected and pre-processed, as shown in Fig. 1. After pre- processing and eliminating duplicates, 4800 reviews were furthered to processing, with positive, negative, and neutral reviews.

The distribution of the class was balanced to train the model and obtain the classification results. It was a good idea to verify the prior chances of unbiased that lead to unfair class distribution. We received 3600, 816, 336 and 48 reviews for each sentiment, totaling 4800 reviews. Subsequently, we eliminated redundant data from the reviews, such as the URL, punctuation, and special characters, because these elements do not offer valuable data in sentiment classification. The proposed architecture is shown in Fig.2 and briefly explained step by step using the pseudocode used for review. The pre-processing is shown in procedure-1, feature extraction is demonstrated in procedure 2 and 3.

Procedure 1: Steps in pre-processing:

Input: Review data obtained from Review server for major hospitals

Output: Pre-processed Reviews

Parameters:

Rg : Regular Expression in python library , Rw : Review, Gw: Get word, Tg: Tag, Tgd: Tag Data, Prw : Parse Review

Tg = Rg Tgd = {

Rg .1[Adjectives] Rg .2[Nown] Rg.3[Verb] Rg.4[Adverbs]

}

Prw = Tgd. get [Tg[0]]

If Prw

return Prw else

return 3

Procedure 2: Steps in Tokenization using lemmatizer

Input: Parsed Review data obtained from Pre-process Review server for major hospitals

Output: Tokenized Reviews

Parameters:

Rg : Regular Expression in python library , Rw : Review , Prw : Parse Review ,CRW : column Review Ps : Porter stemmer , w = word

Initialization: Prw <-[NULL]

Begin Rw in CR do: X = load[Rw] Prw <- X[Rw]

End

For each Rw in Prw do

Rw = Rg .sub['[a-z,A-Z]' for each CRW[Rw[i]] Rw = Rg .lowercase[Rw]

Rw = Rg.split[Rw]

End

Create main root stem for each word in Rw

Rw =[Ps .stem[w] for each Rw [w] if not w in set(stopwords.w('English') Rw = ' ' , join Rw

End

Procedure 3: Creating Bag of words using corpus

Input: Tokenized Review data obtained from Tokenization Output: cleaned Review word ready for model Parameters:

CR : Corpus,, Rw : Review

Initialization:

CR <-[NULL]

For I in range(0, Rw) If(i+1)%100==0):

CR . append [Rw]

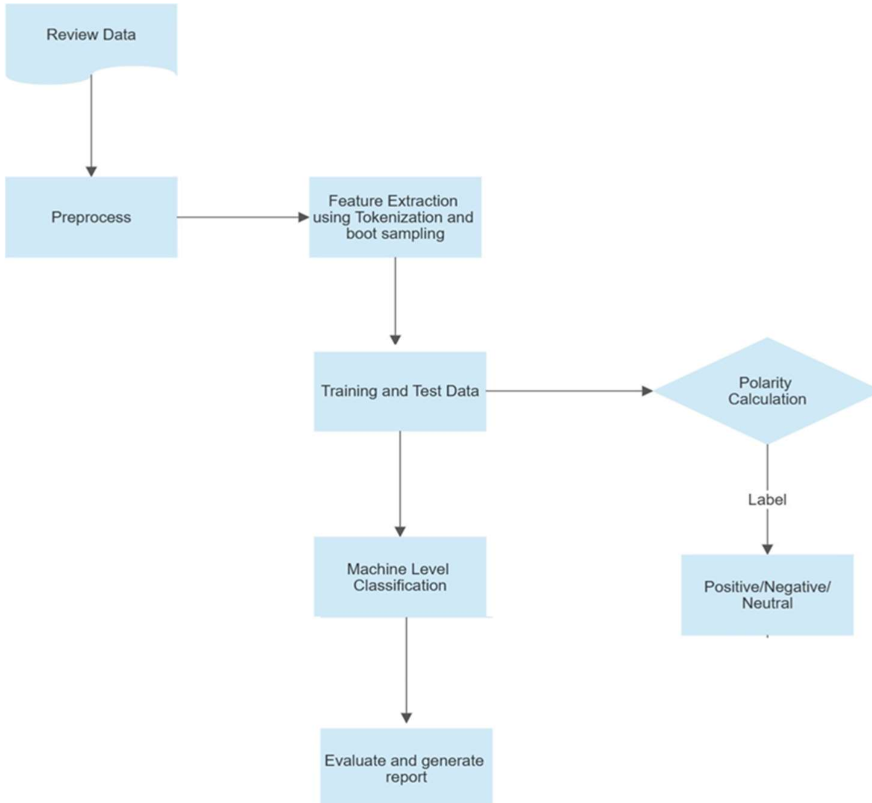

Fig. 2 Overview of Proposed Architecture

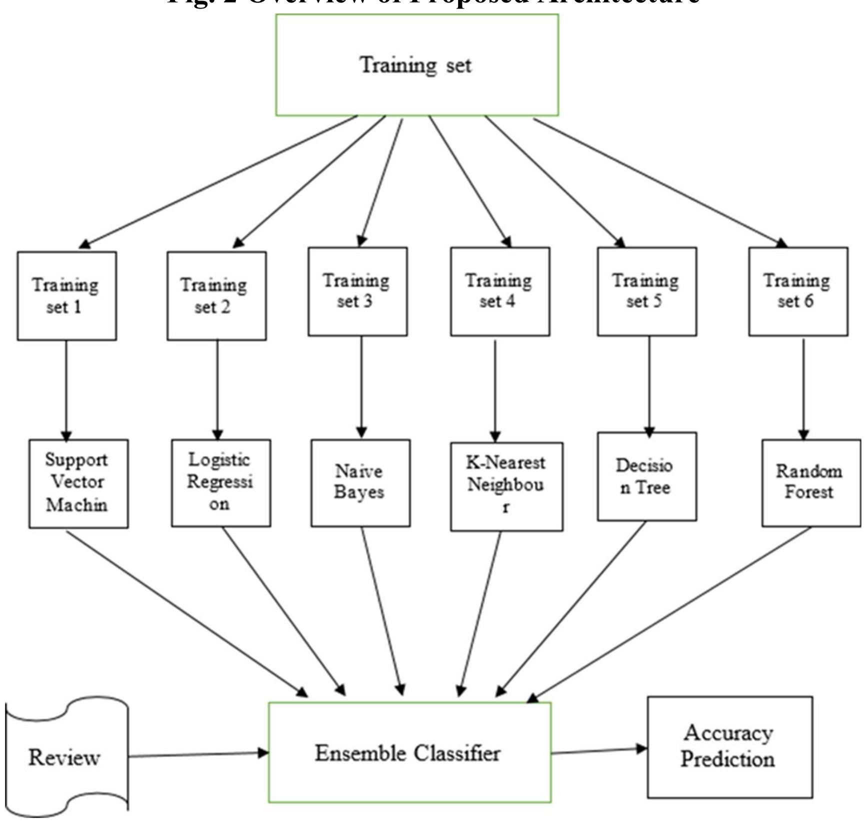

Fig. 3. SCSP Ensemble model

After pre-processing the data, feature extraction is used to tokenize the data and create bag of words using Corpus. As expected, the texts are not in formal English in social reviews such as correct spelling, grammar, and so on. To overcome this problem, a procedure called stemming was executed, where the numerous repetitions of words stemmed together. For instance, "likes" and "liked" can both be stemmed from the same stem word "like." It also aids in the reduction of feature sparsity.

Unigrams, bigrams, and other features can be used in text sentiment classifiers. However, the idea behind considering N-grams is that they can provide more information on various aspects of sentiment than unigrams and bigrams. N-gram, Corpus Approach, stop words, and lemmatization are used for representing weighted featured vectors. The N-gram approach supports review analysis because it tenders to a string of words within a fixed window size 'n'. Pseudo code used for review tokenization/lemmatizing and bag words is shown in pseudo code procedure-2 and 3.
Using the procedure 3, articles have been removed with Bag-of-words in python using stop words extraction. The detailed explanation of the above procedure 2 and 3 is explained as follows:

2.3 PROPOSED CLASSIFICATION APPRAOCHES

After the pre-process, functions have been extracted, and the machine learning classifiers are implemented for sentiment analysis. After carefully analyzing the benefits and detriments of various studies, the following methods were chosen for the present work. SVM , LR , NB, RF, DT and KNN.

2.3.1 Support-Vector-Machine (SVM)

SVM is a popular method to analyze emotions where the binary classification problem is divided into two classes, and the data points of classes are distinguished using a cutoff threshold. Multiple thresholds may appear, but the best one that separates the maximum margin classes will be considered. This model has been used in our study as it has a fair success rate in sentiment analysis.

2.3.2 Logistic Regression (LR)

LR solves classification problems wherever the result is always 0 or 1. In our study, the threshold was set with a probability of 0.5.

2.3.3 Naïve Bayes (NB)

NB also popularly well-known model in sentiment analysis. Bayes rule is applied, which can be represented as:
$P(X|Y) = P(Y)P(Y|X)\ P(X)$ (1)
We have used this model by creating the frequency of word vector, and eventually, the model tested.

2.3.4 K-Nearest Neighbour (KNN)

KNN classification algorithm is popular in sentiment analysis for recognizing a pattern. The attribute values are normalized from Euclidien distance by overweighing attributes with a minimum range. In machine learning classification of sentiment analysis, it is a widely used approach that gave an accuracy of 86% in previous studies.

2.3.5 Random Forest (RF)

RF approach gives a classification tree for binary outcome and the probability estimate for class members. It uses bagging, where the original sample replaces the sample of features. The model is tuned to Scikit -learn library where the number of estimators was 400 and the maximum tree depth was 200. This study tests the performance of XGBoost decision tree model. However, resultant does not show variation, thus not included.

2.3.6 Multi-classification – Ensemble Model

The study employs seven classification models. Cross validation with 10-fold validation has been used to gauge ML classifiers SVM, LR, NB, RF, DT and KNN. The DT had the less accuracy compared to the others, and the bagging classifier had the same accuracy as the RF, therefore, both were removed from the comparison. To predict the accuracy of the overall training dataset, the proposed ensemble model collaborates with the six sentiment classifiers. It was split as divisions and trained independently, as shown in Fig.3.

The SCSP model uses the majority voting method for Bootstrap sampling for classifying the reviews. This process considers the six classifiers with the same weights. While each rating is classed distinctly by these classifiers, the ultimate result predicts the outcome for each class that receives the most votes in the six classifiers. Procedure 4 depicts the majority voting pseudocode.

Procedure 4: Creating voting using Majority voting method

Input: Processed Review data obtained after processing
Output: Predicted accuracy for each model
Parameters:
CRw : classification results for Review, SR: Sentiment Classes Result, CRwi = class1, CRwj = class2 CRwk = class 3
Initialization:
For i in SR and CRwi , CRwj, CRwk in CRw CRwi > 3
then class = CRwi Else if CRwj = 3 then class = CRwj Else if CRw k < 3 then
class = CRwk return class

As described in procedure 4, reviewed data is allocated with a class when six classification results cannot be predicted by the majority voting method. For example, suppose a review is classified to positive by SVM, negative in LR, positive by RF, neutral by DT, positive by KNN, and negative by NB. In that case, that review is allocated to any one of the classes in three classes. In our study we used 10 -fold cross validation, same dataset, to train the six classification models individually. The model process is trained, and model classification is achieved using Python 3.6 version. To recommend the hospital, the sentiment polarity is calculated and categorized into positive, negative and neutral based on cosine similarity amongst reviews.

3. RESULTS AND DISCUSSION

All The performance of the data set is exposed to sentiment analysis, which is classified as positive, negative, or neutral. Generally, over-fitting occurs when we train a small data set; to surpass this challenge, we used cross-validation in our study. It is a model validation method in which samples are split into two data subsets, one for model training and the other for model validation. Because we used 10-fold validation, the data subsets are divided into ten. Nine data subsets were used to train the model, with the residual one used to validate the model. The average is the overall result of this validation method. Testing is done on both a 5-class and a 3-class data set.

The 5-class data set includes outstanding, positive, neutral, negative, and poor ratings, whereas the 3-class data set only includes positive, neutral, and negative ratings. The 3- class is created by combining outstanding with positive and poor with negative from the 5- class. The model's accuracy was validated using several metrics, including precision, recall, and F-measure. Cross validation of accuracy metrics of various machine learning models was used to select the best model performance. The classification report of each sentiment classifier is shown in Table1, it compares various classifier models using cross validation before ensemble and class binding methods. Results shown in Table 1 revealed that Random Forest provide high accuracy of 86.32% which is best out of all other classifiers. Hence this is best suited for reviewing healthcare sentiments

Table 1: Model classification before ensemble

Accuracy of SVM: 0.8333			
	Precision	Recall	F1- Score
	0.7776	0.7778	0.7778
Accuracy of Logistic Reg: 0.7500			
	0.7143	0.5556	0.6250
Accuracy of Naïve Bayes: 0.7059			
	0.7500	0.6667	0.7059
Accuracy of KNN: 0.8021			
	0.7353	0.8525	0.7246
Accuracy of Random Forest : 0.8632			
	0.8000	0.8235	0.8116

CONCLUSION

In this paper, a novel Ensemble Model to Analyze Patient Health Care Opinion Systems. The Classification models are used to classify patients' feelings as positive, negative, or neutral using a machine learning approach to predict superlative models in data analysis. Four different classifiers are used for classification purpose later ensemble method is implemented based on majority of voting. Among the classifiers used random forest classifier performs well in contributing the voting process with better accuracy. While KNN classifier is better in recall value. A prediction obtained with the novel model will chose decent hospital based on the polarity of the review.

REFERENCES

[1] Levinson W et.al.,(2007), “Disclosing medical errors to patients: A status report in 2007, 177, pp.265–267

[2] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, Vol.5, Iss.1,pp.1-167.

[3] M. Abdul-Mageed et.al.(2014), “Subjectivity and sentiment analysis for Arabic social media”, Comput. Speech Lang 28 , pp.20–37.

[4] M.T. Khan et.al.,(2016), “Sentiment analysis for health care, in Big Data: Concepts, Methodologies, Tools, and Applications”, pp. 676–689

[5] Mehta S.(2015) ,” Patient Satisfaction Reporting and Its Implications for Patient Care. The AMA Journal of Ethic. Vol.17,Iss.7, pp.616-621.

[6] Michael Kuhn et.al., (2010), A side effect resource to capture phenotypic effects of drugs, MOL. SYST. BIOL. Vol.6 ,Iss.1 ,pp. 343.

[7] Abdelminaam, D. S., Neggaz, N., Gomaa, I. A. E., Ismail, F. H., & Elsawy, A. A. (2021). Arabicdialects: An efficient framework for Arabic dialects opinion mining on twitter using optimized deep neural networks. Ieee Access, 9, 97079-97099.

[8] Rameshbhai, Chaudhary Jashubhai, and Joy Paulose. "Opinion mining on newspaper headlines using SVM and NLP." International Journal of Electrical and Computer Engineering (IJECE) 9, no. 3 (2019): 2152-2163.

[9] Abdelminaam, D. S., Neggaz, N., Gomaa, I. A. E., Ismail, F. H., & Elsawy, A. A. (2021). Arabicdialects: An efficient framework for Arabic dialects opinion mining on twitter using optimized deep neural networks. Ieee Access, 9, 97079-97099.

[10] Wisnu, Hilman, Muhammad Afif, and Yova Ruldevyani. "Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes." Journal of Physics: Conference Series. Vol. 1444. No. 1. IOP Publishing, 2020.

[11] von Wedel, P., Hagist, C., Liebe, J. D., Esdar, M., Hübner, U., & Pross, C. (2022). Effects of hospital digitization on clinical outcomes and patient satisfaction: nationwide multiple regression analysis across German hospitals. Journal of Medical Internet Research, 24(11), e40124.