Automated Feedback System for Customer Reviews using Sentiment Analysis

S.V.Sudha[1]*, S. Sankar Ganesh[2] , S.Ayyasamy[3], Anush Kumar N[4], Burhanuddin M[4], Mohana Priyan K[4]

Professor, Department of Computer Science and Engineering, Dr.N.G.P. Institute of Technology, Coimbatore, India.
Professor, Department of Information Technology, Dr.N.G.P. Institute of Technology, Coimbatore, India.
Professor,Department of Computer Science and Enginerring,Dr.N.G.P.Institute of Technology, Coimbatore , India
Under Graduate Scholars, Department of Computer Science and Engineering, Dr.N.G.P.Institute of Technology, Coimbatore, India.
*Corresponding author Email: svsudha.mvenki@gmail.com

ABSTRACT

Today, public reviews in terms of digital nature play a vital role in customer buying pattern and customer satisfaction. People may not be able to read all public/customer reviews of a product/movie and get the relevant feedback. The aim of the paper is to automate the process for providing effective feedback to customers by extracting subjective information from the text or feedbacks. This approach is unique in the following way: in case of positive feedbacks the customer is greeted with an appreciation message else a message conveying that services would be improved is sent to customers. In this proposed research, a model has been developed which performs sentiment analysis by using Natural Language Tool Kit (NLTK). The classification was done using Naive Bayes classifier by calculating the probability of the reviews of the customers with the highest value is considered as either positive or negative. As a result, the emotions and sentiment analysis improves psychology of the customer, reduces the stress and improves product purchase rate.

Key words : Sentiment Analysis, Naive Bayes, Polarity, Natural Language Processing.

1. INTRODUCTION

The current affordable and ubiquitous generation of Web provides substantial amount of opinionated social big data which facilitates decision making. Sentiment Analysis has gained Multimedia Tools and Applications importance as quick as a wink for tracking the mood/view of people by analyzing this unstructured, multimodal, informal, high-dimensional and noisy social data. It helps in gaining insights 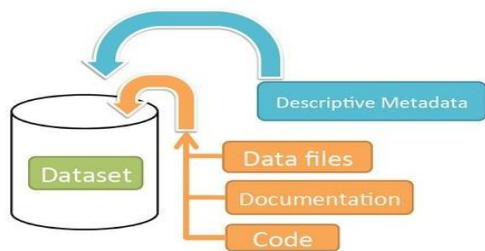for a particular subject, topic, event or a matter within various domains such as market, business or government intelligence. Various studies in this field are [1] proposes a state-of-the-art research for aspect-based sentiment analysis of Arabic Hotels reviews using two implementations of long short-term memory (LSTM) neural networks. [2] huge volume of such information has to be technically processed for segregating the relevant knowledge. Sentiment analysis is the popular method extensively used for this purpose. [3] Current sentiment analysis work focuses on improving granularity at aspect level serving two distinct goals such as aspect extraction and feeling classification of product reviews and meaning classification of target-dependent deep learning tweets. [4] rapid increase in the number of applications and total app store revenue has accelerated app store data mining and opinion aggregation studies. While developing companies and app store regulators have pursued upfront opinion mining studies for business intelligence and marketing purposes, research interest into app ecosystem and user reviews is relatively new. [5] Natural Language Processing techniques for identifying features of fine-grained apps in the reviews. The work then extract the user 's feelings about the areas selected and give them a general score across all reviews. [6] models of neural networks also incorporating user behavioral information within a document (tweet). A Convolutionary Neural Network (CNN), is the neural network used in this paper. [7] Aims to predict the actual emotions from the raw text expressed by the people through the field of natural language processing (NLP). The main aim of our work is to process the raw sentence from the Twitter dataset and find the actual polarity of the message. [8] constructed a joint architecture which places of RNN at first for capturing long-term dependencies with CNNs using global average pooling layer while on top a word embedding

method using GloVe procured by unsupervised learning in the light of substantial twitter corpora to deal with this problem. [9] Multimodal sentiment analysis model to determine the polarity of feeling and the score for any incoming tweet, i.e. textual, image or infographic and typographic. [10] consumer sentiment is generally positive towards several famous brands. Analyzing brands tweets using both a qualitative and quantitative approach. This study brings intensity and depth to the discussion about attitudes to cosmopolitan brands. This paper is made up of diverse modules.[16]

1.1 Data Set

A data set (or dataset) is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds to a given record of the data set in question. The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set. Each value is known as a datum. Data sets can also consist of a collection of documents or files as referred in Figure 1.
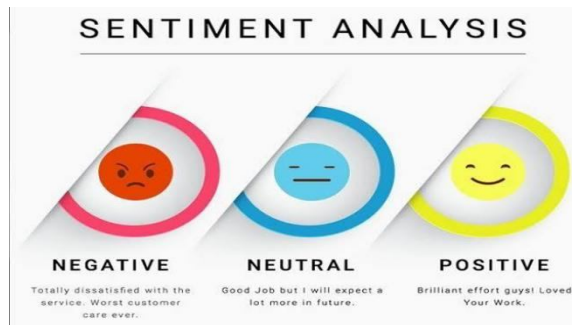

Figure 1: Schematic diagram of a dataset.

1.2 Sentiment Analysis

Sentiment Analysis is a field within Natural Language Processing (NLP) that builds systems that try to identify and extract opinions within text. These systems usually extract attributes of the expression, besides identifying the opinion, as shown in Figure 2.[12][13]

Figure 2: Attributes of Sentiment Analysis

• Polarity: when the speaker expresses a positive or negative opinion.
• Subject: that which is being discussed.
• Holder of opinion: person or entity expressing an opinion..

1.3 Emails

The collected reviews are split based on its opinion (positive, negative or neutral) and a feedback mail is sent to the respective users individually. The email is sent based on the customer sentiment being portrayed within the review.

2. MATERIALS AND METHODS

The content of feelings may be about books, people, hotels, products, research, events, etc. These feelings are becoming very beneficial to corporations, governments and individuals. Whilst the content is to be useful, a large part of the content generated requires the use of text mining techniques and the analysis of feelings. But the process of sentiment analysis and evaluation confronts several challenges. These challenges become obstacles in analyzing the exact meaning of feelings and detecting the appropriate polarity of sentiment. Sentiment analysis is the practice of using natural language processing techniques and text analysis to identify and extract subjective information from text. Sentiment Analysis is basically classified into two comparisons:[11][10]

First comparison discusses the relation between the challenges of the sentiment analysis and the structure of the review.

Second analysis explores the importance of overcoming the problems of sensing in order to enhance accuracy.

2.1 First comparison is between the thirty-seven papers on the research. The aim of this comparison is to recognize the relationship between the challenges of feeling and the structure of review and how to impact on the results of feeling. The structure of sentiment review becomes an important factor which affects the selection of major challenges (referred in Figure 3)

2.1.1 Structured sentiments are found in formal sentiment reviews but are targeted as books or research on formal issues.

2.1.2 Semi-structured feelings lie within the range of formal structured feelings and unstructured feelings. These demand understanding of several reviews issues. This form, which depends on Pros and Cons, is described separately by the writer and typically short sentences are the contents of Pros and Cons.[17]

2.1.3 Unstructured Sentiments are an informal, free format of text. Pros and Cons are not formally separated, and the content may consist of several sentences, where each sentence contains features and/or opinions. For the example below, unstructured reviews have the potential to provide more detailed information about opinions than

their counterpart Explicit feature, where a feature appears in a review sentence segment / chunk, that feature is called an explicit product feature. For instance; the image is wonderful in the chapter, picture is an explicit function. Implicit function, if a function f does not appear in the analysis section but is inferred, then an implicit feature of a company is named. For instance; it's very expensive in the section, the price is an implied feature and expensive is an indicator of features.[14][15]
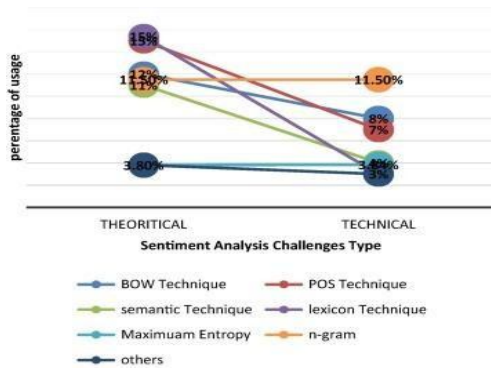
Figure 3: SA challenges and technique types

2.2 Second comparison explains the summary of the challenges of feeling and how to improve each one's accuracy based on previous works. Its aim is to identify the most significant sentimental challenges and how to improve their pertinent results for the techniques used. Illustration. 3 Explains the proposal to use the techniques in relation to the challenge types of sentiment analysis (SA) (Theoretical or technical). The theoretical challenges use many techniques to improve the results with solving the selective sentiment challenges.

Each and every customer and their review are being analyzed based on the review and categorized into the polarity of the reviews that is whether it is positive, negative review. The categorizing of the reviews based on the polarity is carried out by making use of the Naïve Bayes algorithm. The process consists of collecting a data set to analyze the reviews based on the sentiment portrayed within the review. The data set is collected by the process of web scraping and the reviews are being analyzed using Sentiment Analysis algorithm. After analyzing, the reviews are being stored into the database. It is represented by the below Figure 4.

The next module is the Short Message Sevice(SMS) module, where the mobile number of the customer is linked with the email id. Once linked an SMS is also being sent to the customer based on the sentiment of the review. The SMS module is inherited using "Twilio" API for sending SMS using python.

2.3 Naïve Bayes

Naïve Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with strong independence assumptions between the features. They are among the simplest Bayesian network models. Naïve Bayes is one of the supervised classification techniques which classify the text/sentence that belongs to particular class. It is the probabilistic algorithm which calculates the probability of each word in the text/sentence and the word with the highest probability is considered as output.

Example:

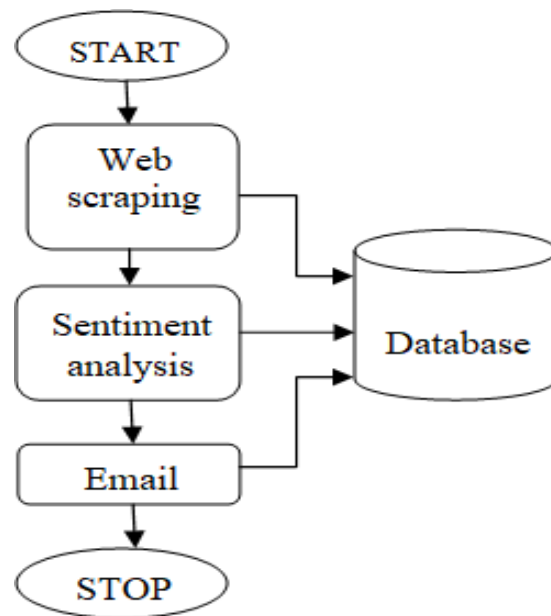"It is a beautiful day today."

"Today I will go outside"

"I am going to go for a movie"

Now tokenization is performed on these lines then get Line 1: It, is, a, beautiful, day, today.

Line 2: Today, I, will, go, outside.

Line 3: I, am, going, to, go, for, a, movie.

Figure 4:Block Diagram for Sentiment Analysis for collection of data

Now need have to maintain a table which contains 2 columns with attributes word and counts i.e., the module is going to count frequency of each word in the document.

TABLE 1: Word Count Frequency

Word	Count
It	1
Is	1
To	1
a	2
Rain	1
Today	2
I	2
Go	2
Beautiful	1
Outside	1
Movie	1
For	1
Will	1
Day	1

Now after completion of the table, the words that frequently appeared are noted and considering only the words that are fixed in different documents. As of the above Table 1, here are a total of 14 words but after that, there are only 4 most frequent words which reduce the comparison.

The additional feature that is being imposed in this system is the sending of response mails to the customers those who have dropped the reviews. The response is an automated mailing system that sends mail to the customer based on the review of the respective customer. An appreciation mail is sent to the customer if it is an positive review (i.e.)"We are pleased that you were satisfied by the service provided by us and look forward for your presence." and an apologizing mail is sent to the customer if it is an negative review (i.e.) "The product quality will be improved for further customer satisfaction."

2.4 Twilio

Twilio is an organization based in San Francisco, California, cloud networking platform as a service (CPaaS). Twilio allows software developers to make and receive phone calls programmatically, send and receive text messages and use its web service APIs to perform other communication functions.

The Twilio API for python is being integrated into the python environment and is being used to send sms. The procedure for registering and sending sms is as follows:

- Log in to Twilio and get your first Twilio phone number enabled by SMS

- Build your development environment for sending outbound messages

- Give your original outbound text message

- Establish your Twilio Messaging Service first

- Send a second SMS to that Service

The Work have incoporated the twilio API into our module to allow sending programmed sms to the customer along with an email. The SMS module consists of feeding the customers mobile number linked with the email id and an sms is being sent to the customer based on the sentiment analyser review (ref in Figure 5)

Figure 5: SMS sent using twilio API

Both appreciation and apologizing sms are being sent to the customer as are being sent via email. This module is incoporated because many people do not check their email and so sms will help the customer in getting a feedback for their review and it will increase customer satisfaction.

3 RESULT AND ANALYSIS

Figure 6: Sentiment Analysis of reviews from an Amazon Data Set.

Figure 7: Reviews from various websites with unprofessional English and Smiley.

Figure 8: Sentiment Analysis of Multiple Line Reviews.

In Figure.6 an Amazon dataset for movie reviews is being used and the reviews are being trained and tested to identify the sentiment. In Figure.7 reviews are being collected from various websites with unprofessional English and emojis .Then these reviews have been trained and tested to obtain the sentiment portrayed within the review. In Figure.8 Reviews are being collected consisting of multiple lines while testing each and every line within the review is being tested to identify the sentiment and obtain the probability. In conclusion, the proposed work have developed a model which performs sentiment analysis by using Natural Language Tool Kit (NLTK). The classification was done using Naïve Bayes classifier by calculating the probability of the reviews of the customers with the highest value is considered as either positive or negative. This model can further be enhanced to any desired level if one wants to by incorporating more features in the database.

REFERENCES

1. M. Al-Smadi, B. Talafha, M.Al-Ayyoub and Y. Jararweh **Using long short-term memory deep neural networks for aspect based sentiment analysis** in *Arabic reviews International Journal of Machine Learning and Cybernetics*, 2018, Vol 10,pp 2163-2175

2. A.Kumar and G. Garg. **Systematic literature review on context-based sentiment analysis** in *social multimedia Multimedia Tools and Applications* , 2019, Vol 79, pp 15349-15380

3. F.Aisopos, G.Papadakis, K.Tserpes and T.Varvarigou **Content vs. context for sentiment analysis: a comparative analysis over microblogs** in *Proceedings of the 23rd ACM conference on Hypertext and social media*, 2012 pp. 187–196.

4. S.Deng, A.P.Sinha, and H.Zhao Resolving **ambiguity in sentiment classification: the role of dependency features.** *ACM Trans Manag Inf Syst (TMIS)*, 2017,pp 8(2–3):4

5. Frankenstein, K.Joseph and K.M.Carley **Contextual sentiment analysis**. in: Xu K, Reitter D, Lee D, Osgood N (eds) *Social, cultural, and behavioral modeling. SBP-BRiMS 2016. Lecture notes in computer science*, Springer, vol 9708.

6. H. Dohaiha. H Prasad, P Maag, and A. Alsadoon, **A Deep Learning for Aspect-Based EmotionAnalysis:** *A Comparative Study Expert Applications Systems*, 2019 Vol 118, pp 272-299

7. N.Genc-Nayebi and A.Abran **A systematic literature review** in *Opinion mining studies from mobile app store user*, 2017, Vol 125, pp 207–219

8. M.Arora and V. Kansal **Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis** in *Social Network Analysis and Mining*, 2019. Vol 12

9. F. Abid, M.Alam, M.Yasir and C.**Li Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter** in *Future Generation Computer Systems*, 2019 Vol 95. pp 292-308

10. M.Mohamed and **Mostafa More than words: Text mining for consumer brand feelings** in *social networks. Application-expert systems*, 2013, Vol 40(10), pp 4241–4251

11 . S.Sankar Ganesh, S.V.Sudha, S.Ayyaswamy, G.Pandiya Rajan and S.Chidambaram, **Optimal feature subset selection and msvm classification based cbir for medical images** in *International Journal of Advanced Trends in Computer Science and Engineering, 2020,Vol* 9(4), pp. 4746-4752

12. Rudy Prabowo and Mike Thelwall **Sentiment analysis: A combined approach** in *Journal of Informetrics,* 2009, Vol 3, pp 143– 157.

13. Muhammad Zubair Asghar, Aurangzeb Khan, Shakeel Ahmad and Fazal Masud Kundi**, A Review in Sentiment Analysis on Feature Extraction i**n *Journal of Basic and Applied Scientific Research*, 2014, Vol 4(3),pp 181-186.

14. N.V.S Pavan Kumar, J.K.R Sastry, K Raja Sekhara Rao **Mining Negative Frequent regular Itemsets from Data Streams** in in *International Journal of Emerging Trends in Engineering Research* Vol 7, No.8 August 2019, pp 85-98

15. S.V.Sudha, M. Rathidevi, R. Yaminipriya **Trends of cryptography stepping from ancient to modern** *in International IEEE Conference on Innovations in Green Energy and Healthcare Technologies*,2017, pp 1-9.

16. Rostam Sadiqi, Mohd Zainuri Saringat, Aida Mustapha, Salama and A MostafaSequentia**l Recursive Structure Searching Algorithms for Arabic and English Texts** in *International Journal of Emerging Trends in Engineering Research,* 2020,Vol 8,No 1.1, pp 62-67

17. N.Indurkhya, F.L.Damereau**, Handbook of Natural Language Processing.** 2nd Ed. 2010.