Call Centre Sentiment Analysis

[#1]Prof. Pankaj Khambre, [#2]Mr. Vivekanand Patil, [#3]Ms. Swapnali Pawar,
[#4]Ms. Prajakta Bobade, [#5]Mr. Rohit Badgujar

[1]vivekanandpatil21@gmail.com,
[2]swapnalip77@gmail.com,
[3]prajaktasbobade@gmail.com,
[4]badgujarrohit37@gmail.com

[#1]Assistant Professor, Department of Information Technology,

[#2345]U.G. Student, Department of Information Technology,

G.H. Raisoni institute of Engineering, Pune, India.

ABSTRACT

A sentiment analyser is used to analyse opinions, feedback of customer; categorize it into positive and negative. The input uploaded by the administrator in the form of mp3 is pre-processed. After uploading the audio file, conversion to text is performed. The words in the sentence which are of no use are removed (stop words) post which stemming is performed. Important words which are useful are counted using count vector. Naïve Bayes algorithm is implemented to classify if it is positive or negative. Accuracy is calculated. Negation handling is also performed which analyses complex statements.

Keywords: Sentiment analysis, call center data analysis, classification

ARTICLE INFO

I. INTRODUCTION

In BPO, customers provide feedback in the form of audio. The amount of data is huge. Appointing a person to analyze data is tedious task and it is important to analyze this data, hence a need for automating this task is crucial. This will help company's analyze the service provided by them, the market value of product. By analyzing feedback, it will help the company serve customer better and provide customer satisfaction. That is why are developing and implementing a system, we can automate this process, by converting audio to text and providing analysis if it is positive and negative. Also, the aspect of each entity can be analyzed. This will automate the process by converting the audio to text and provide an analysis in less time. The proposed system will be helpful in business scenarios where the sentiment of the customer can be analyzed to predict the future of product or polarity of feedback in order to evaluate the service provided by BPO. Usually, sentiment analysis runs on NLP (Natural Language Processing) which assists machine to understand the text and looks for specific keywords. The sentences are also analyzed in terms of aspect where a sentence is broken into parts namely entity, attribute. Post which a bar graph represents the analysis of sentiment based on entity. For sentiment analysis, the sentence is broken into entity, attribute, and polarity of the feedback. Hence the user will be provided with a simple bar chart for easy understanding of analysis.

II. LITERATURE SURVEY

[1] In this paper, Opinion mining is tracking people' s opinion posted on web for products, politics, organizations, persons, events, issues and analysing them. It helps the manufacturers what the customers want and if consumers have a positive or negative feedback or requesting some enhancements. Aspect extraction for sentences is also performed to check what aspect of product is being analysed by customer by performing stemming, stop words removal and analyzing them.

[2] In this paper, Feedback provided by customer is analyzed for BPO. The purpose of this paper is to identify the intention of the caller and attitude of the customer. To determine if they are satisfied with the product or the changes requested are addressed and successfully completed.

[3] In this paper, scope of sentiment analysis is defined. Four levels have been identified on which analysis can be done. Level 1 consists of analysis of sentence which can be review of a product. The analysis can be positive, negative, neutral. Level 2 considers the entire document as a single unit. The whole document is considered for analysis of positive, negative or neutral. Level 3 is based on aspect. Each sentence consists of entity and attribute. Analysis of each sentence based on attributes/ aspect of entity is categorized Into good, bad, neutral. Level 4 consists of analysis of relationship between two users using graph theory which can be implemented for recommendation system.

[4] In this paper, Sentiments expressed by users on Twitter are analysed using Machine learning algorithms. It performs detailed analysis on tweets using multinomial logistic, support vector regression. On twitter the restriction for users is that post can contain maximum of 280 characters, slangs, abbreviations, irregular expressions .The classification of sentiment analysis is also done using Decision tree & Random forest. For testing purpose twitter data has been made publicly accessible. Here the opinion of public expressed on brand, politics, event is classified in positive, negative, neutral.

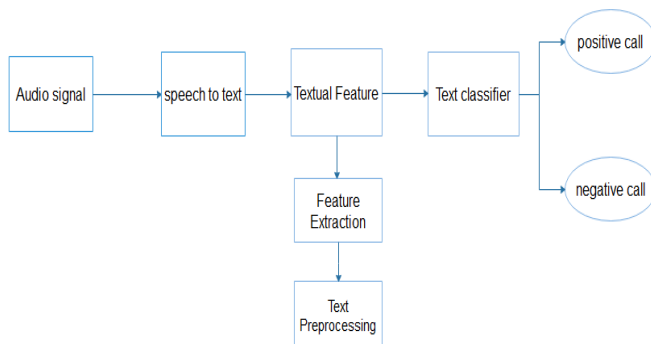## III. PROPOSED SYSTEM

Stanford's POS Tagger

Fig 1. Proposed architecture

The input uploaded by the administrator in the form of mp3 is pre-processed. After uploading the audio file, conversion to text is performed.
Stop words are removed and stemming is performed in order to get the root words. Important words which are useful are counted using count vector. Naïve Bayes algorithm is implemented to classify if it is positive or negative.

Algorithm:

The call recordings are converted into text form using IBM Watson speech API

1. Input: The feedback provided by customer is uploaded in mp3 format.

2. Audio to text conversion: The call recordings are converted into text form using IBM Watson speech API.

3. Pre-processing: The converted text is then preprocessed by filtering out stop words and by performing stemming.

4. Aspect based Sentiment analysis:
Aspect regarding all the attributes are then mapped and classified under positive or negative class label.

5. Bar graph generation:
By counting the positive and negative aspects of any attributes, bar graph is generated.

6. Product suggestion:

The user can select features with top most priority and the system suggest the product by performing analysis

IV. MATHEMATICAL MODEL

Let S be a system that describes the execution of the application.
S = {I,O,M, S,F} , Identify the modules as M, S={M,...}

M={E,R}
where,
E = Predefined Questions .
R = Undefined Questions.
System Description: Input: fAudio recordings in MP3 formatg
Output: fY1, Y2g
Where, fY1, Y2g
belong to Success
Functions: fFpreprocess(),   FgraphGenerate(), FsuggestProduct()g
Fpreprocess()= fFk1(), Fk2()g
Fk1()= stemming()
Where,
Input ip= fset of words or sentencesg
Output op= froot wordsg
If ip doesnt belong to the set of prede_ned words, then skip those words
else convert it into its root form.
Fk2() = removestopwords()Where;
Input ip = fset of words or sentencesg
Output op= fsentences after removing stopwordsg
If ip belongs to set of prede_ned words like is, the etc, then remove
them
FgraphGenerate():
Where, Input ip = fAspect wise polarityg
Output op = fAspect1 (+,-), Aspect2 (+,-)....g
FsuggestProduct(): Input ip= f2 aspects choseng
Output op= fComparative analysis performed to suggest correct productg
Success: fY1, Y2g
Where, Y1 = fGraph generatedg
Y2 = fProduct suggested to customerg
Failure:fNULLg

Above mathematical model is NP- Complete

V. CONCLUSION

The feedback provided to BPO is not being analysed properly and relies heavily on manual analysis. It becomes a tedious task to analyse it manually hence automating this task is done by sentiment analyser. It also provides result in the form of bar chart to help understand the review at a glance along with each aspect of the feedback.
Currently the system is developed for English language and does not support multiple languages. Hence there is a great scope for adding regional languages to the dictionary so language will not be a barrier.

VI. REFERENCES

[1] Extracting Aspects and Mining Opinions in Product Reviews using Supervised Learning Algorithm, IEEE SPONSORED 2ND INTERNATIONAL CONFERENCE ON ELECTRONICS AND COMMUNICATION SYSTEMS (ICECS '2015)

[2] Customer Voice Sensor: A Comprehensive Opinion Mining System for Call Center Conversation, 2016 IEEE International Conference on Cloud Computing and Big Data Analysis.

[3] Sentiment Analysis Over Social Networks: An Overview, 2015 International Conference On Advanced Informatics: Concepts, Theory And Application ICAICTA).

[4] Twitter Sentiment Analysis Based on Ordinal Regression, 2019 5th IIAI International Congress on Advanced Applied Information.