ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 7, July 2022

Sentiment Analysis on Student Feedback Using RNN

[1] Ayisha Sumayya K, [2] Roshni Joseph, [3] Saisree R K, [4] Sneha Satheesh T, [5] Deepthi P, [6] Rahna C M

[1] [2] [3] [4] Student, Department of Computer Science and Engineering,
KMCT Collage of Engineering for Women, Kerala, India.
[5] [6] Teacher, Department of Computer Science and Engineering,
KMCT Collage of Engineering for Women, Kerala, India.
Corresponding Author Email: [2] roshnijosepht@gmail.com

Abstract— ***Almost all organizations collect feedback to improve the quality of service provided. Frequent and ongoing feedback analysis is inevitable for the development of any organization. But handling this huge amount of data is time-consuming and not economical. In the present feedback system, qualitative analysis of data is not possible. It may require huge manpower to analyze the feedback. Also, the currently used grading system is not useful in conveying the exact emotions of students. To overcome this, a system is proposed that gives summarised feedback and works on sentiment analysis. The summarised feedback contains deep key phrases. Key phrases can provide highly condensed and valuable information that helps in acquiring the main idea of the content. It attempts to capture the deep semantic meaning of text content. Here, the Recurrent Neural Network (RNN) algorithm is used for keyphrase extraction and sentiment analysis.***

Index Terms— ***Feedback Summarising, Keyword Extraction, Machine learning, Recurrent Neural Network (RNN), Sentiment Analysis***

I. INTRODUCTION

Collecting feedback frequently is a significant way to improve the achievement in learning. Through feedback collection the emotion of student for each teacher can be understood. Effective feedback, both positive and negative are valuable information that will be used to make important decisions. Educational institutions are one of the significant areas that require feedback for the improvement of the existing learning process, teaching material, and evaluations. But handling these qualitative opinions is a challenging task. The process is timeconsuming and not economical as it requires manpower.

In most cases, the main purpose of feedback is not utilized. So, we propose a software that can analyze the sentiment from the feedback and detect keywords based on their nature. The frequency of emotion is plotted as graph. High-quality keyphrases allow the understanding, organizing, and accessing of the content. This predicts the aspect described and the orientation of the given feedback. The basic idea of the keyphrase generation model is to compress the content of source text into a representation. After training, key phrases are generated and we get the predicted word sequences with the highest number of probabilities.

In the existing system, the qualitative feedback is collected manually or using a grading system. A grading system cannot be considered a good approach as the actual feedback is not evaluated. By this method exact emotion of the student cannot be identified in the case of manually collected feedback, the data is huge. So, the chance of completely avoiding the feedback is high. If the whole feedback needs to be considered, then there arises the problem of hiring people for feedback evaluation, which is not economical. And if there is any correction in the given feedback then it is difficult to correct or even not possible sometimes. In the existing system, the feedback is collected by a teacher or any other staff. The disadvantages of existing system are: The lack of security of the data, requires more manpower, it is more time consuming, it consumes a large volume of paperwork, it requires manual calculations, and it has no direct role for higher officials, and overall performance analysis is not easy. The proposed system overcomes these problems of the conventional method.

The deep key phrase generation helps in acquiring the main idea of text content without losing what the students have conveyed. Sentiment analysis and key phrase extraction is done using RNN (Recurrent Neural Network) algorithm. Finally, a graph is generated that gives the overall opinion of students about each faculty.

Contributions: The contributions can be:

- We introduced a feedback system to collect feedback from the students, the students can upload feedback for each teacher. The feedback can be uploaded in the form of both text and voice
- We implemented a speech-to-text conversion system that covers the feedback given in the form of a speech-to-text format.
- We propose a generative model for keyphrase extraction based on sequence to sequence learning which enables the model to successfully predict both

IFERP connecting engineers... developing research

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 7, July 2022

present and absent phrases, as well as out-of-vocabulary words; we use the yake library for keyword extraction.

- The system also predicts emotion from given feedback. The feedback is analyzed processed and emotion is predicted, whether it is a positive, negative or neutral emotion.
- We also generate a graph based on the emotions predicted. The percentage of emotions of the teacher is analyzed and then based on that value the graph is plotted.

The paper flows as: We introduce the literature survey in Section II and then justify the intuitions of our method in Section III. Next, we introduce our experiment setups in Section IV and the results are given in Section V. Finally, our findings are concluded in Section VI.

II. LITERATURE REVIEW

In this paper sentiment analysis is performed using different algorithms in different platforms, algorithms like Naive Bayes,

SVM, ANN, and RNN are analyzed. With the help of the analysis made the accuracy of the algorithms is determined and the best algorithm is chosen for the implementation of the system. Naive Bayes algorithm shows up to 80% accuracy, SVM shows up to 85% accuracy, ANN shows up to 87% and RNN shows more than 88% accuracy. RNN is used in this system as it shows more accuracy.

Sentiment analysis using the Naive Bayes algorithm is used for analyzing the positive and negative opinions in the collected data. These factors can then help e-commerce companies focus on improving service and company quality which will be associated with increased traffic, sales, and company profits. Then data collection process, data cleaning, and lexicon classification is performed. R Studio is used to process the three stages. R Studio is a software application that uses the basis of the R programming language. Here the data collected are Twitter tweets. A data crawler is used to collect the tweets. The Naive Bayes method is used here to classify the sentiment levels. There are mainly 4 steps they are collecting data, text parsing, tokenization, and text mining using the naive Bayes algorithm.

- Collecting data: In this data collection process, data tweets obtained using crawler data from Twitter are used.
- Text parsing: Parse the tweets by describing them verbatim is text parsing.
- Tokenization: Tokenization is the process of processing sentences into several words that have been separated from characters and taken into words that have value. That is cleaning the tweets and selecting meaningful words.
- Text mining: Text mining is performed by using the Naive Bayes method. We calculate the value of the class probability by dividing the number of class data by the total or number of existing documents.

Up to 3000 data tweets are analyzed and about 80% accuracy is achieved [1].

Recognition of emotions is a problematic issue in the analysis of natural languages since a lot of data is stored. Various methods were developed and applied to recognizing emotions in multiple applications such as image, voice, video, and text. Also, emotional understanding in applications where interaction occurs between humans and machinery will increase robots' performance. However, in photographs and videos, feelings are computational, really costly, and challenging to identify. Proposes a method for emotion detection from text using hypergraph-based SVM. The HISVM displays various emotional relationships. The new scheme is introduced in the field of text emotion recognition. Here seven simple emotion categories such as anger, fear, disgust, sadness, guilt, and surprise are used for textual emotion recognition. A newly trained classifier recognizes feeling from text input by applying a learning model. Human emotion can be recognized from the text by classifying text based on some emotional classes.

For that, a dataset is prepared for each emotional class. Data Collection is found and then classify to test cases and training scenarios. Word is tokenized and text was converted to the appropriate sequence. We use the word embedding technique to classify positive and negative emotion-based words. Using this, different types of emotion can be detected. Using data types precision rate and accuracy rate on types of emotions were found. The Hypergraph-based support vector machine is used for hypergraph regularisation and to build our baseline since it's a classification-based work. That form of SVM provides the desired accuracy. [2].

Increase in access to the internet more people are joining and participating in social media services, which led to the generation of a large amount of data on social media platforms like Facebook, Youtube, Twitter, etc. This paper performs sentiment analysis for 'iPhone X' using Twitter API, Facebook API, and data collected from news websites using python as a platform. ANN is applied to the collected data using R programming and removes unwanted words, special characters, and spaces. the dataset may include the date, number of positive tweets, number of negative tweets, number of neutral tweets, and polarity. The polarity and sentiment of the tweet were recorded using the inbuilt library TextBlob. The data collected is partitioned into train and test datasets. The training data is used to train the neural network, it uses a back-propagation algorithm. And test dataset is used to determine the accuracy of the result, it uses a feed-forward neural network. After training the neural network a neural net graph was plotted and the output of all the neural net was the polarity of the input data [3] [5].

Explains the generation of deep key phrases using RNN. In

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 7, July 2022

this paper, RNN encoder-decoder model is used. Here the source content is compressed to a hidden representation using an encoder and generates corresponding keywords using the decoder. A keyphrase is a short piece of text that summarizes the main semantic meaning of a longer text. High-quality keywords help to understand, organize, and accessing of document content. As a result, many studies have focused on ways of automatically extracting keyphrases from the textual content. A bidirectional GRU is used as an encoder and a forward GRU as a decoder. The attention mechanism is combined with RNN to automatically locate relevant input components. Furthermore, the coping mechanism is incorporated to help the model in finding important parts based on the positional information. RNN with a copying mechanism could predict the words that are out of vocabulary but in the source text. Thus the model can generate keywords by understanding the content of source text as real annotators, regardless of the presence or absence of a key phrase. In this paper, phrases that do not match any source text are denoted as absent keyphrases. The keyphrases that fully match a part of the text are denoted as present keyphrases. When ranking phrase candidates, these features only target to detect the importance of each word in the document with respect to the statistics of word occurrence and co-occurrence and are unable to reveal the full semantics that underlies the document content. To overcome the drawbacks of previous studies, the process of keyphrase prediction is examined again with a focus on how real human annotators would assign keyphrases. When a document is given, human annotators will firstly read the text to get a basic understanding of the content, then they try to understand its essential content and then summarize it into keywords It represents long text with representative short text chunks [4].

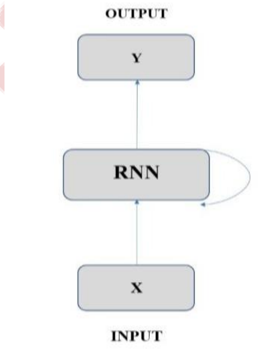


Fig. 1. RNN process

III. PROPOSED METHOD

The system helps in qualitative analysis of the feedback. It analyzes the sentiment from the feedback and detect keywords based on their nature. High quality keyphrases allow the retrieval of content when organized. The emotion detection is performed by opinion classification on the data-set using RNN. There is separate login for student, teacher and admin. The users can log into the system using their credentials. Students can upload the feedback either in text or speech form. Speech is converted to text using API. The collected feedback is then used to generate deep key phrases and to analyze sentiment using RNN. The sentiment analysis is done into main three classes, positive, negative, and neutral.

The system has three modules, admin, teacher and student. Admin is assigned overall management of system and can view feedback about all teachers. The teacher has permission to view feedback about oneself. Students can upload feedback in the form of speech or text. Voice is then converted into text and is stored in the database. The data-set contain feedback and its analysed sentiment.The model of data-set if shown in figure:5. The emotions are classified into good, bad and average.The final result is a graphical representation of the analyzed sentiment. We propose a generative model for keyphrase extraction based on sequence to sequence learning which enables the model to successfully predict both present and absent phrases, as well as out-of-vocabulary words. The feedback is analyzed processed and emotion is predicted, whether it is a positive, negative or neutral emotion. We also generate a graph based on the emotions predicted. The percentage of emotions of the teacher is analyzed and then based on that value the graph is plotted. The deep key phrase generation helps in acquiring the main idea of text content without losing what the students have conveyed.

RNN is specialized for processing a sequence of data, x(t)= x(1), . . . , x() with t is the time step index ranging from 1 to . For the tasks that involve sequential inputs, such as language and speech, it is often better to use RNNs. In an NLP problem, if you want to predict the next word in a sentence we need to know the words before itself. RNNs perform the same task for every element of a sequence, with its output being dependent on the previous computations, so they are called as recurrent RNNs have a internal memory which captures information about what all has been calculated so far. RNN is a class of ANNs (Artificial Neural Network), here the connections that are between the nodes will form a directed or an undirected graph along the temporal sequence. This feature allows the exhibition of the temporal dynamic behaviour. RNN are derived from feedforward neural networks. RNN can make use of their internal memory to process variable-length sequences of inputs. This makes them applicable to tasks such as connected and unsegmented handwriting recognition or speech recognition. Recurrent neural networks are theoretically going to complete and it can also run arbitrary programs to process arbitrary sequences of inputs.

Algorithm Steps:

1. Training data set
2. Model creation using RNN

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 7, July 2022

3. Feature extraction
4. Login to the system
5. Upload feedback
6. Extract deep keyphrases
7. Detect emotion
8. Graph is generated
9. End

In fig 1 it shows the proposed sytem block diagram. The system is trained using a dataset. During training a tokenised pickle file,and a model will be created. The tokenise pickle file contain the tonised dataset vectors. Using the created model prediction is performed. The model is loaded and if the values in that model is 0 the it is neutral, if 1 then negative, otherwise positive.

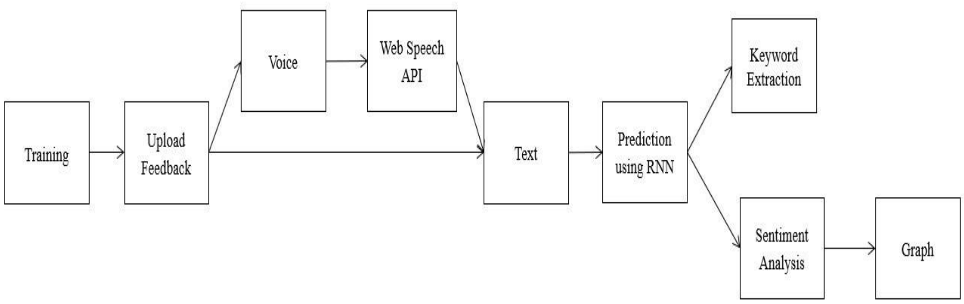


Fig. 2. Block diagram

The labels and text is collected and tokenised. The messages and labels are converted into 80% training and 20% testing. Then tokenised and then a model is created, using this model, The features are trained data is given to the mode. The tokenised data is compressed as pickle file. the data is then accessed from this pickle file. Keras library file is used for prepossessing. Then the features are extracted. The student can login to the system using their credentials and upload feedback for the respective teacher. Using RNN algorithm java the deep key-phrases are extracted and the emotion is predicted. The percentage of emotions is calculated and with respect to that the graph is plotted using json file.

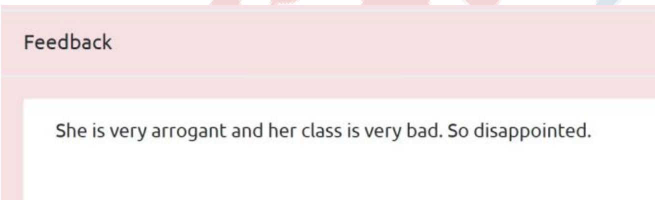


Fig. 3. Upload feedback

Fig. 4. extracted key words of Uploaded feedback

IV. EXPERIMENTAL SETUP

A. Speech to text conversion

For speech to text conversion, we use Web Speech API. The Web Speech API is designed for speech analysis and speech synthesis. It allows web users to send speech input to web applications. The web applications use the Web Speech API to transform the speech into text.

There are good reasons why speech input in web applications might be beneficial for users.

- opening the field for a new input method enhances accessibility. Just like Braille Touch enables visually impaired users to type without looking, speech in- put might provide a convenient means of alternative input.
- speech is a hands-free input method.
- users have become more and more used to speech input on mobile phone applications and might demand that competing web applications have the same capabilities.

For the speech- to- word conversion, we use Web Speech API. The Web Speech API is designed for speech analysis and speech conflation. It allows web people to transfer speech input to web uses. The web operations use the Web Speech API to transfigure the speech into word.

The Web Speech supports the conversion of speech to textbook and vice versa. The Web Speech API allows people to record audio from the microphone, which is also transferred via an HTTPS POST request to the speech recognition web service. The result can also be reused within the limits of a JavaScript operation running in the Chromium surfer. The TSP Speech Database serves as a gold standard in this evaluation. It's one of the countable annotated speech recognition data sets that are available on the world wide web at no cost.

The database contains greater than 1400 utterances from 12 men and 12 women speakers. In total, there are 720 lines distributed over 72 lists of 10 line each. This list is known as the Harvard lines. The recordings were performed in an auditory anechoic room and designed to have a low noise position. Fig 4 shows the dataset. All rulings are in English, and utmost of the speakers were adult native speakers. Each line in the TSP Speech Database is independently transferred to the speech recognition web service. It appears that the web

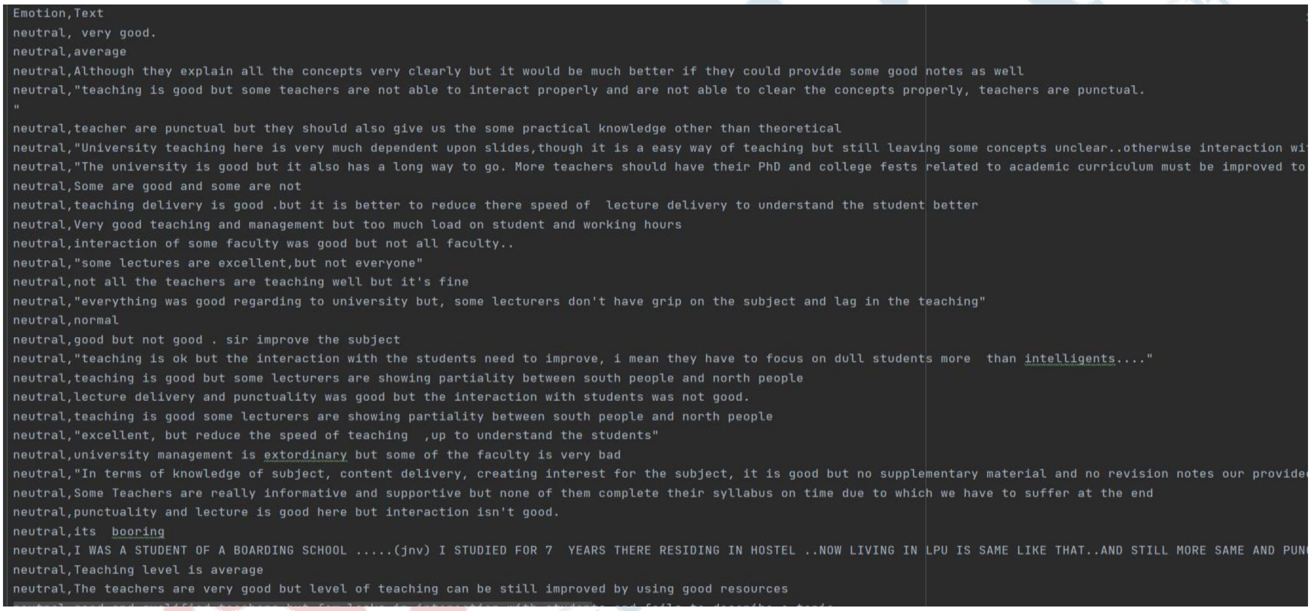
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 9, Issue 7, July 2022

service accepts audio data in the FLAC format. Conversion from the surge- encoding of the dataset to the FLAC encrypting needed by the web service is loss-less, which is fortunate for evaluation. Differences in case, whitespace, and punctuation aren't taken into account.

The Web Speech API is event- grounded, which fits by well with the rather call-back-heavy style of typing code with JavaScript. The event- grounded framework allows programs to asynchronously use speech. Events are also used to report intermediate speech recognition results which are suitable because it allows programs to give nearly immediate feedback to the people.

B. Key-phrase extraction

Student uploads feedback for separate instructor. The collected feedback is analysed and keywords are taken. the taken keyword represent a summary of the collected feedback Still, they're substantially grounded on the statistical parcels of the words and do not really take into account the semantic aspects of the full document. KeyBERT is a minimum and easy- to use keyword taking fashion that aims at working this issue. Keyphrase provides a piece of largely- condensed information that can be effectively used for understanding, organizing, and reacquiring word content. YAKE! is a lightweight unsupervised automatic keyword taking system which rests on words statistical features taken from single documents to choose the most important keywords of a sentence, which builds upon original sentence statistical features taken from single documents; i.e., it doesn't need any training corpus [13].



```
Emotion,Text
neutral, very good.
neutral,average
neutral,Although they explain all the concepts very clearly but it would be much better if they could provide some good notes as well
neutral,"teaching is good but some teachers are not able to interact properly and are not able to clear the concepts properly, teachers are punctual.
"
neutral,teacher are punctual but they should also give us the some practical knowledge other than theoretical
neutral,"University teaching here is very much dependent upon slides,though it is a easy way of teaching but still leaving some concepts unclear..otherwise interaction wi
neutral,"The university is good but it also has a long way to go. More teachers should have their PhD and college fests related to academic curriculum must be improved to
neutral,Some are good and some are not
neutral,teaching delivery is good .but it is better to reduce there speed of  lecture delivery to understand the student better
neutral,Very good teaching and management but too much load on student and working hours
neutral,interaction of some faculty was good but not all faculty..
neutral,"some lectures are excellent,but not everyone"
neutral,not all the teachers are teaching well but it's fine
neutral,"everything was good regarding to university but, some lecturers don't have grip on the subject and lag in the teaching"
neutral,normal
neutral,good but not good . sir improve the subject
neutral,"teaching is ok but the interaction with the students need to improve, i mean they have to focus on dull students more  than intelligents...."
neutral,teaching is good but some lecturers are showing partiality between south people and north people
neutral,lecture delivery and punctuality was good but the interaction with students was not good.
neutral,teaching is good some lecturers are showing partiality between south people and north people
neutral,"excellent, but reduce the speed of teaching  ,up to understand the students"
neutral,university management is extordinary but some of the faculty is very bad
neutral,"In terms of knowledge of subject, content delivery, creating interest for the subject, it is good but no supplementary material and no revision notes our provide
neutral,Some Teachers are really informative and supportive but none of them complete their syllabus on time due to which we have to suffer at the end
neutral,punctuality and lecture is good here but interaction isn't good.
neutral,its  booring
neutral,I WAS A STUDENT OF A BOARDING SCHOOL .....(jnv) I STUDIED FOR 7  YEARS THERE RESIDING IN HOSTEL ..NOW LIVING IN LPU IS SAME LIKE THAT..AND STILL MORE SAME AND PUN
neutral,Teaching level is average
neutral,The teachers are very good but level of teaching can be still improved by using good resources
```

Fig. 5. Data-set

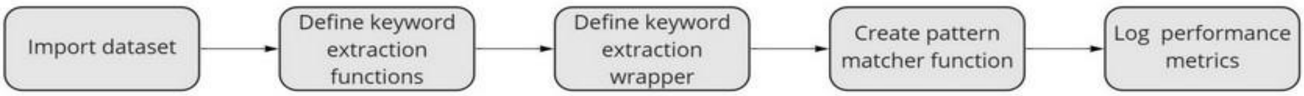


Fig. 6. Steps in keyword extraction

The input document is embed using a pre-trained BERT model. Keywords and expressions (n- grams) are retrieved from the same document. Each keyword is also put together into a fixed- size vector with the same model used to embed the document. The steps in keyword extraction is shown in fig 6.

In fig 3 the feedback uploaded is shown and fig 4 shows the extracted keywords. The very first import the dataset that has the collected textual data. We 'll also produce separate functions that apply the retrieval reason. further we 'll produce a function that applies the retrieval on the entire corpus. Spacy will also help us define a matcher object that will return true or false if a keyword matches a syntactic pattern that makes sense for our task. Eventually we 'll pack up everything in a function that produce our final report. Each extractor takes in as an argument the line from which we want to retrieve keywords and returns a list of keywords, from the suitable to the worse according to their importing fashion.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 7, July 2022

C. Emotion detection

The system predicts emotion from given feedback using RNN algorithm. For accurate prediction, the dataset used here is developed by collecting feedback and its emotion from students of different schools and colleges. On training the dataset, a pickle file is generated. This enables to make new predictions at a later time, without having to train the model again. Pickling converts python object structure into a byte stream. The feedback is analyzed processed and emotion is predicted, whether it is a positive, negative or neutral emotion [7] [8].

V. RESULT

The system predicts the emotions after analysing the feedback and the keywords are extracted.

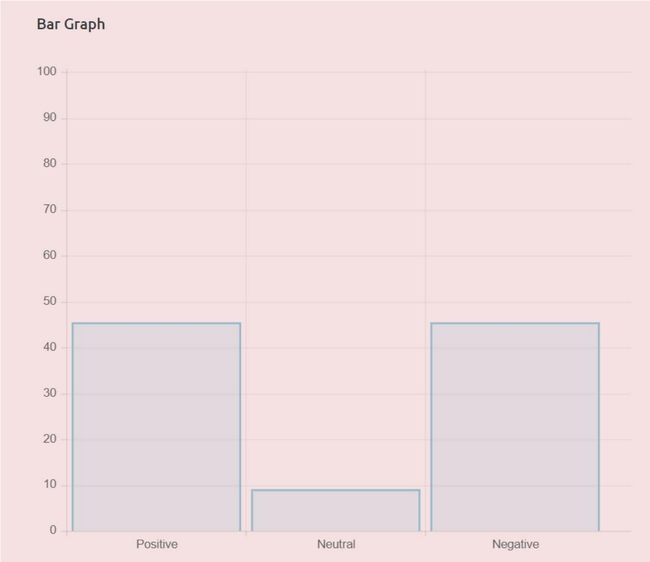


Fig. 7. Graph generated

There is about 90% accuracy in the system. With the rate of predicted emotions graph is plotted. Fig 6 shows how the graph is viewed for admin and teacher, where 60% of predicted emotion is positive and 40% of emotion predicted is negative[9].

VI. CONCLUSION

This paper uses a supervised learning technique, opinion mining to detect the emotion of the student. This student feedback system collects feedback from the students. Sentiment analysis is performed on the collected feedback to detect the emotion of the student. The generated emotions are plotted into a graph in which it is possible to understand the overall rating of the performance of a particular teacher. This strategy will be useful to analyze which teaching method can be the most beneficial for students, which in turn helps improve their achievements in learning. Automating the student's feedback may give several advantages together with saving price, time, and creating economic report generation, etc. The utilization of opining mining will facilitate summarizing the feedback report effectively and evaluating performance.

This system can be implemented in any sector due to its wide range scope. In the future, this system can be designed in such a way that not only the English language but other native languages can also be used. Also, the input speech can be directly used for sentiment analysis which is faster and more precise[10]. This will help in understanding the emotion from the depth, tone, phonetics, etc of speech.

VII. ACKNOWLEDGEMENT

We would like to thank the faculty of Department of Computer Science and Engineering, KMCT College of Engineering for Women for the supporting and guiding us in the research and implementation of the software.

REFERENCES

[1] Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler : Twitter, Meylan Wongkar; Apriandy Angdresey
[2] Hypergraph Regularized SVM and Its Application Emotion Detection ,Md Helal Hossen, Wenjun Hu
[3] Sentiment Analysis and Prediction using Neural Network, Sneh Paliwal, Sunil Kumar Khatri and Mayank Sharma
[4] Deep keyphrase Generation, Rui Meng, Sanqiang Zhao, Shuguang Han, Daqing He, Peter Brusilovsky, Yu Chi
[5] Comparative Study of CNN and RNN for Natural Language Processing, Wenpeng Yin, Katharina Kann, Mo Yu and Hinrich Schutze
[6] Real-time Convolutional Neural Networks for Emotion and Gender classification, Octavio Arriaga, Matias Valdenegro-Toro, Paul Ploger¨
[7] Motion classification using deep neural networks and emotional patches, Jungming Huang, Xiangmin Xu, Tong Zhang
[8] Emotion Detection and Recognition from Text Using Deep Learning, Chew-Yean
[9] Evaluating and Summarizing Student's Feedback Using Opinion Mining, Vaibhav Jain
[10] Speech Emotion Recognition: Methods and Cases Study, Leila Kerkeni1,
[11] Youssef Serrestou1, Mohamed Mbarki, Kosai Raoof1 and Mohamed Ali [11] Web Speech API, Julius Adorf
[12] A Text Feature Based Automatic Keyword Extraction Method for Single
[13] Documents, Ricardo Campos, V´ıtor Mangaravite
[14] KeyBERT: Keyword Extraction using BERT, Prakhar Mishra
[15] Introduction to Recurrent Neural Network, https://www.geeksforgeeks.org/introduction-to-recurrent-neuralnetwork/
[16] Keyword Extraction — A Benchmark of 7 Algorithms in Python,Andrea D'Agostino, https://towardsdatascience.com/keyword-extraction-abenchmark-of-7-algorithms-in-python-8a905326d93f