SENTIMENT DETECTION IN NATURALISTIC AUDIO USING NLP

[1]Roopa B S, [2]Aparna
[1]Assistant Professor, [2] Assistant Professor,
[12]Electronics and Communication Engineering,
[12]J N N College of Engineering, Shimoga, Karnataka, India

Abstract: Emotions and sentiment of software developers can largely influence the software productivity and quality. However, existing work on emotion mining and sentiment analysis is still in the early stage in software engineering in terms of accuracy, the sizeof datasets used and the specificity of the analysis. In this work, we are concerned with conducting entity-level sentiment analysis. Sentiment analysis and opinion mining is the field of study that analyzing peoples opinions, sentiments, evaluations, attitudes, and emotions from written language. It is the most active research areas in natural language processing and is also widely studied in data mining, web mining, and text mining. In fact, this research has spread outside of computer science to the management sciences and social sciences due to its importance to business and society as a whole. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forum discussions, blogs, micro-blogs, Twitter and social networks. For the first time in human history, we now have a huge volume of opinions recorded in digital form for analysis. The framework illustrates the different facets of analysis to be considered whileperforming sentiment analysis and, hence, serves as a new benchmark for future research in this emerging field.

Index Terms – Sentiment, accuracy, mining, natural language processing, benchmark.

I. INTRODUCTION

Quantifying users content, idea, belief, and opinion is known as sentiment analysis. User’s online post, blogs, tweets, feedback of product helps business people to the target audienceand innovate in products and services. Sentiment analysis helps in understanding people in a better and more accurate way. It is not only limited to marketing, but it can also be utilized in politics, research, and security.

There are mainly two approaches for performing sentiment analysis.

- Lexicon-based: count number of positive and negative words in given text and thelarger count will be the sentiment of text.
- Machine learning based approach: Develop a classification model, which is trainedusing the pre-labelled dataset of positive, negative, and neutral.

Human communication just not limited to words; it is more than words. Sentiments are combination words, tone, and writing style. As a data analyst, it is more important to understand our sentiment. Sentiment analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials.Generally speaking, sentiment analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. The attitude may be his or her judgement or evaluation affective state, or the intended emotional communication. Sentiment analysis is the process of detecting a piece of writing for positive, negative, or neutral feelings bound to it. Humans have the innate ability to determine sentiment; however, this process is time consuming, inconsistent and costly in a business context. It's just not realistic to have people individually read tens of thousands of user customer reviews and score them for sentiment.

Keyword spotting is the naivest approach and probably also the most popular because of its accessibility and economy. Text is classified into effect categories based on the presence of fairly unambiguous affect words like happy, sad, afraid, and bored. The weaknesses of this approach lie in two areas: poor recognition of affect when negation is involved and reliance on surface features. About its first weakness, while the approach can correctly classify the sentence today was a happy day as being happy, it is likely to fail on a sentencelike today wasn't a happy day at all About its second weakness, the approach relies on the presence of obvious affect words that are only surface features of the prose .

In practice, a lot of sentences convey affect through underlying meaning rather than affect adjectives. For example, the text my husband just filed for divorce and he wants to take custody of my children away from me certainly evokes strong emotions, but uses no affect keywords and therefore, cannot be classified using a keyword spotting approach[1]. Lexical affinity is slightly more sophisticated than keyword spotting as, rather than simply detecting obvious affect words, it assigns arbitrary words a probabilistic affinity for a particular emotion. For example, accident might be assigned a 75% probability of being indicating a negative affect, as in car accident or hurt by accident, these probabilities are usually trained from linguistic corpora. Though often outperforming pure keyword spotting, there are two main problems with the approach First, lexical affinity, operating solely on the word-level, can easily be tricked by sentences like I avoided an accident (negation) and I met my girlfriend by accident (other word senses) Second, lexical affinity probabilities are often biased toward text of a particular genre, dictated by the source of the linguistic corpora, this makes it difficult to develop a reusable, domain-independent model[2].

Statistical methods, such as Bayesian inference and support vector machines, have been popular for affect classification of texts. By feeding a machine learning algorithm a large training corpus of affectively annotated texts, it is possible for the system to not only learnthe affective valence of affect keywords (as in the keyword spotting approach), but also totake into account the valence of other arbitrary keywords (like lexical affinity) punctuationand word co-occurrence frequencies.

However, traditional statistical methods are generally semantically weak, meaning that, with the exception of obvious affect keywords, other lexical or co-occurrence elementsin a statistical model have little predictive value individually. As a result, statistical text classifiers only work with acceptable accuracy when given a sufficiently large text input. So, while these methods may be able to affectively classify users text paragraph-level, they do not work well on smaller text units such as sentences or clauses.

Sentiment analysis, also called opinion mining, is the field of study that analyzes peoples opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. It represents a large problem space. There are also many names and slightly different tasks, sentiment analysis, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. However, they are now all under the umbrella of sentiment analysis[3]. While in industry, the term sentiment analysis is more commonly used, in academia both sentiment analysis and opinion mining are frequently employed. Regardless, they basically represent the same field of study.

II PROBLEM STATEMENT

From a research point of view, this abstraction gives us a statement of the problem and enables us to see a rich set of inter-related sub-problems which make up the sentiment analysis problem. It is often said that if we cannot structure a problem, we probablydo not understand the problem. The objective of the definitions is thus to abstract a structure from the complex and intimidating unstructured natural language text. They also serve as a common framework to unify various existing research directions, and to enable researchers to design more robust and accurate solution techniques by exploiting the inter-relationships of the sub-problems. From a practical application point of view, thedefinitions let practitioners see what sub-problems need to be solved in a practical system,how they are related, and what output should be produced.

Unlike factual information, opinions and sentiments have an important characteristic, namely, they are subjective. It is thus important to examine a collection of opinions from many people rather than only a single opinion from one person because such an opinion represents only the subjective view of that single person, which is usually not sufficient for application. Due to a large collection of opinions on the Web, some form of summary of opinions is needed.

III METHODOLOGY

A single keyword spotting system (KWS) is developed for sentiment detection. In the new architecture, the text-based sentiment classifier is utilized to automatically determine the most powerful sentiment-bearing terms, which is then used as the term list for KWS.

In order to obtain a compact yet powerful term list, a new method is proposed to reducetext-based sentiment classifier model complexity while maintaining good classification accuracy. Finally, the term list information is utilized to build a more focused language model for the speech recognition system. The result is a single integrated solution which is focused on vocabulary that directly impacts classification.

In the first step towards feature extraction, we extract words and word combinations which potentially depict sentiment. This is done by parsing the raw text using a part-of-speech (POS) tagger. We use POS tagger system. After tagging, words and word combinations formed by particular combinations of adjective (JJ), Verb (V*), Adverb (RB*) and Noun (N) are extracted as sentiment bearing features. Our approach towards sentiment extraction uses two main systems, namely, Automatic Speech Recognition (ASR) system and text- based sentiment extraction system[4]. For text based sentiment extraction, we propose a new method that uses POS (part-of-speech) tagging to extract text features and Maximum Entropy modeling to predict the polarity of the sentiments (positive or negative or neutral)using the text features. An important feature of our method is the ability to identify the individual contributions of the text features towards sentiment estimation. This provides us with the capability of identifying key words/phrases within the video that carry important information. By indexing these key words/phrases, retrieval systems can enhance the ability of users to search for relevant information.

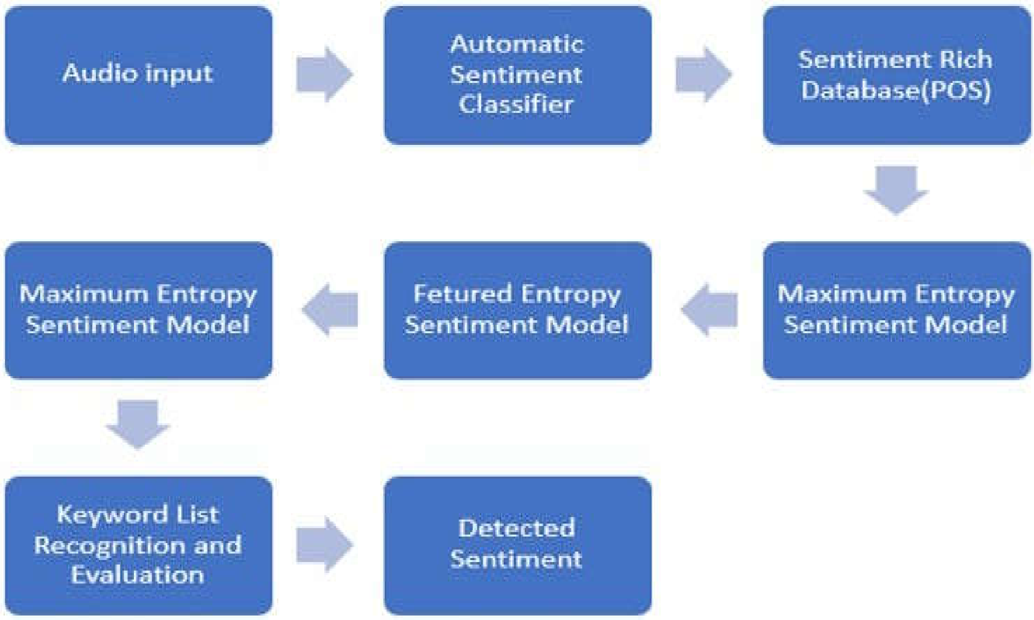

Figure 1.1: Methodology of the proposed system

IV DESIGN AND IMPLEMENTATION

Audio sentiment analysis using automatic speech recognition is an emerging research areawhere opinion or sentiment exhibited by a speaker is detected from natural audio. It is relatively under-explored when compared to text based sentiment detection. Extracting speaker sentiment from natural audio sources is a challenging problem. Generic methods for sentiment extraction generally use transcripts from a speech recognition system, and process the transcript using text-based sentiment classifiers. In this study, we show that this baseline system is sub-optimal for audio sentiment extraction. Alternatively, new architecture using keyword spotting (KWS) is proposed for sentiment detection. In the newarchitecture, a text-based sentiment classifier is utilized to automatically determine the most useful and discriminative sentiment-bearing keyword terms, which are then used as a termlist for KWS. In order to obtain a compact yet discriminative sentiment term list, iterative feature optimization for maximum entropy sentiment model is proposed to reduce model complexity while maintaining effective classification accuracy. A new hybrid ME-KWS joint scoring methodology is developed to model both text and audio based parametersin a single integrated formulation. For evaluation, two new databases are developed for audio based sentiment detection, namely, YouTube sentiment database and another newlydeveloped corpus called UT-Opinion Opinion audio archive. These databases contain naturalistic opinionated audio collected in real world conditions. The proposed solutionis evaluated on audio obtained from videos in youtube.com and UT-Opinion corpus. Our experimental results show that the proposed KWS based system significantly outperformsthe traditional ASR architecture in detecting sentiment for challenging practical tasks.

Text based sentiment detection is an established field in natural language processing (NLP). Sentiment analysis / opinion mining, analyzes people’s opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services,organizations, individuals, issues, events, topics, and their attributes. There is an enormous amount of opinionated data in the social media and on the Web in the form of Twitter, Facebook, message boards, blogs, and user forums. The decision making process of people is affected by the opinions formed by a wide range of thought leaders and ordinarypeople over the web. Amazon, Yahoo, Google and various other personalized websites are a significant resource for obtaining opinions concerning products of any kind. Many consumers form their decision to buy a product dependent on feedback from online reviews. This information not only helps ordinary people make decisions, but also provides indicators for companies about the reception of a product, or a political context, to understand the mood of people regarding an ongoing social/ cultural/political/economic issue. Typically, a given text is classified to exhibit positive, negative or neutral. This form of automatic classification has numerous applications such as measuring public opinion/sentiment using Twitter feed, analyzing online product reviews; understand mass social human behavior over a topic, product or an event. Text based reviews form only one of the many ways people can express their sentiment/opinion about products or socialissues. Audio/Video is also a prominent method to express opinions. Millions of videos onYouTube are about products and movie reviews, product un-boxing, political, social issue analysis and opinions on them.

V SYSTEM DESIGN

5.1 Existing System

Many consumers form their decision to buy a product dependent on feedback from online reviews. This information not only helps ordinary people make decisions, but also provides indicators for companies about the reception of a product, or a political context, to understand the mood of people regarding an ongoing social/ cultural/political/economic issue. Typically, a given text is classified to exhibit positive, negative or neutral sentiment. This form of automatic classification has numerous applications such as measuring public opinion/sentiment using Twitter feed, analyzing online product reviews; understand mass social human behavior over a topic, product or an event.

5.2 Proposed System

First one is the offline text based sentiment model generation, the second is the ASR basedsentient detection system forming our baseline system, and finally the third is a proposed system using audio Keyword Spotting (KWS) approach. Each block is explained in detailin subsequent sections. In reality, accurate sentiment detection generally relies on a small fraction of the speech recognition transcript, because sentiment bearing vocabulary tends to be sparse in spoken opinions. In other words, sentiment detection accuracy depends on being able to reliably detect and recognize a much focused vocabulary in the spoken comment audio stream. Therefore, keyword spotting (KWS) technology is expected to bebetter suited for sentiment detection, as opposed to full-transcript ASR.

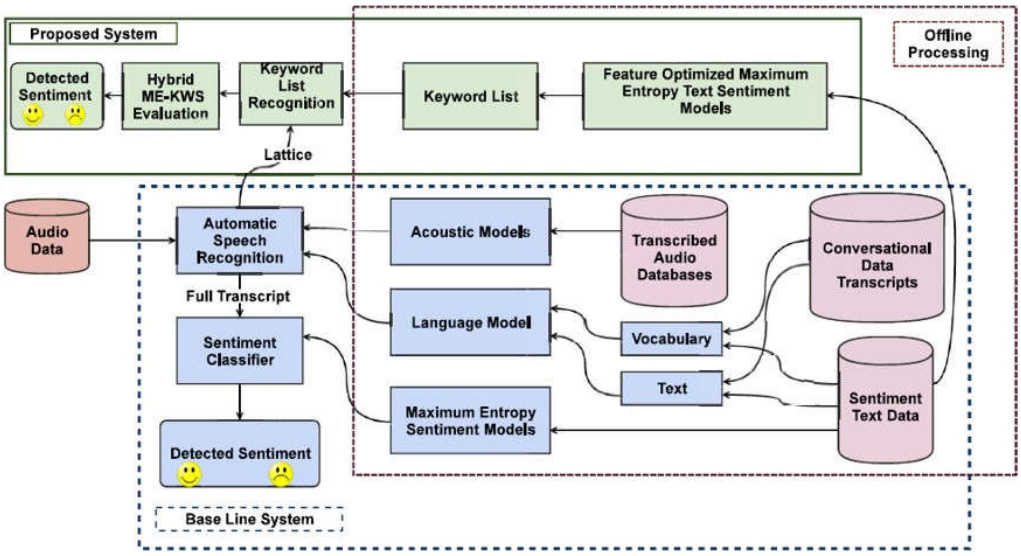

Figure 2: Block diagram of the proposed system

1. A single keyword spotting system (KWS) is developed for sentiment detection. In the new architecture, the text-based sentiment classifier is utilized to automatically determine the most powerful sentiment-bearing terms, which is then used as the termlist for KWS.

2. In order to obtain a compact yet powerful term list, a new method is proposed to

reduce text-based sentiment classifier model complexity while maintaining good classification accuracy. Finally, the term list information is utilized to build a more focused language model for the speech recognition system.

3. The result is a single integrated solution which is focused on vocabulary thatdirectly impacts classification. The proposed solution is evaluated on videos from YouTube.com.

4. In this study, a new system for sentiment detection in audio is presented. While automatic sentiment detection using text is a mature area of research, and significantresearch has been done.

5. On websites such as YouTube, Amazon, Flipkart automatic audio sentiment detection technology would undoubtedly be useful is collecting and summarizing informationfor users.

6. It transforms into a format that the machine can use through natural language processing.

7. It is useful to note that audio sentiment detection concerns with detection of opinion(positive vs. negative), and is different from speech emotion recognition.

8. In the feature extraction, we extract words and word combinations which potentiallydepict sentiment.

9. This is done by parsing the raw text using a part-of-speech (POS) tagger. We use POS tagger system.

10. After tagging, words and word combinations formed by particular combinations of adjective (JJ), Verb (V*), Adverb (RB*) and Noun (N) are extracted as sentiment bearing features.

5.3 Implementation

As shown in Figure 3, initially POS tagging is used to generate a large set of initialtext features that consist of nouns, adjectives, adverbs and verbs. We also extract text features that correspond to adjective-noun, verb-adjective, adverb, adjective and adverb- verb combinations. In the next step, we employ Maximum Entropy (ME) modeling technique to predict the ratings (positive and negative) given the text features extracted from review comments. This constitutes our baseline text-based sentiment model.

Our approach towards sentiment extraction uses two main systems, namely, automatic

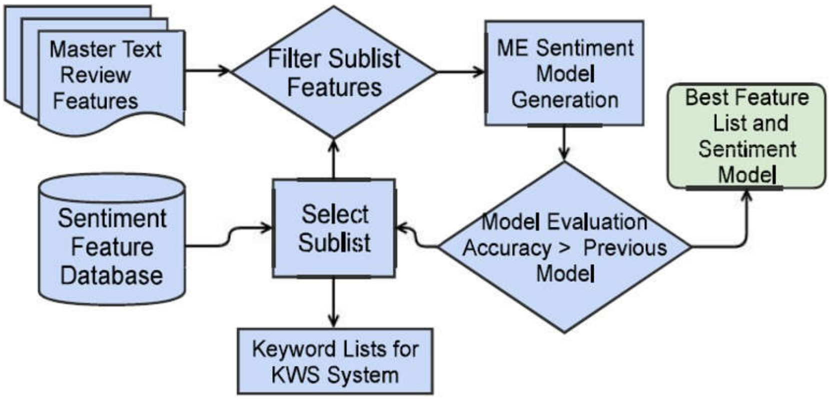

Figure 3: Method of Implementation

speech recognition (ASR) system and text-based sentiment extraction system. For text based sentiment extraction, we propose a new method that uses POS (part-of-speech) tagging to extract text features and Maximum Entropy modeling to predict the polarityof the sentiments (positive or negative) using the text features. An important feature of our method is the ability to identify the individual contributions of the text features towardssentiment estimation. This provides us with the capability of identifying key words/phraseswithin the video that carry important information. By indexing these key words/phrases, retrieval systems can enhance the ability of users to search for relevant information.

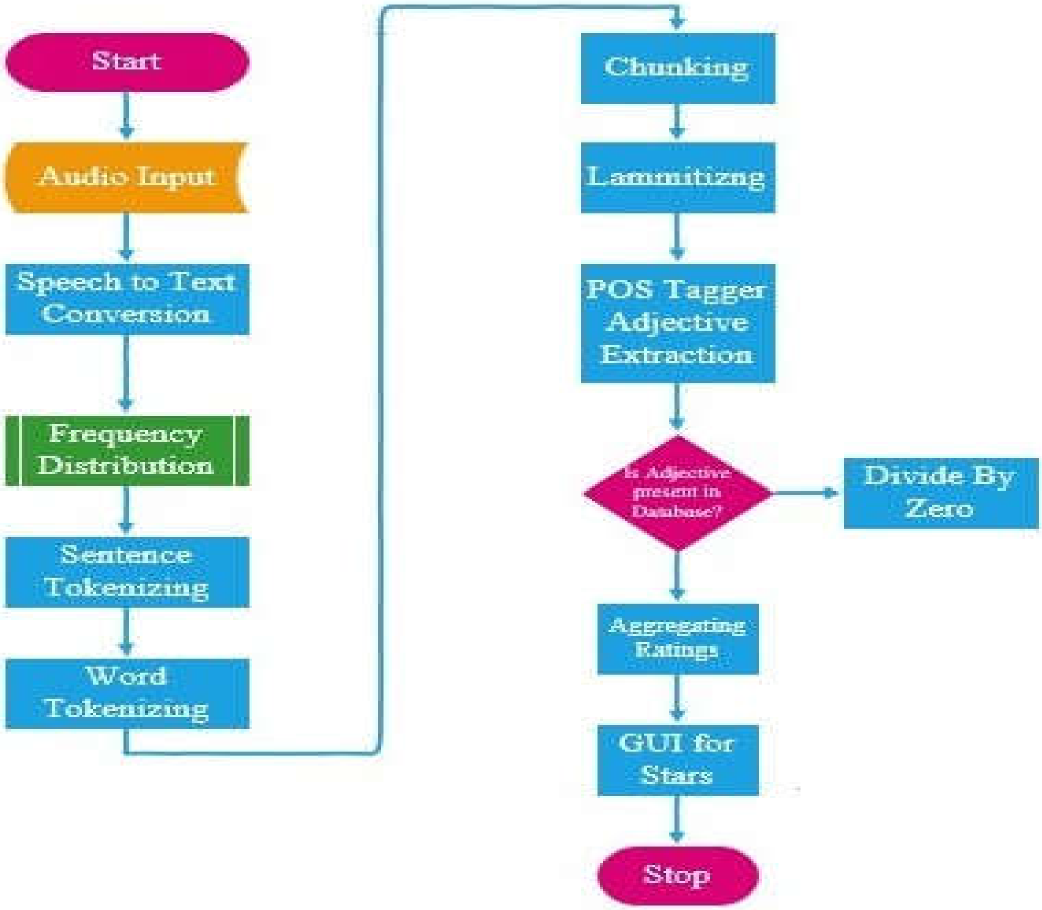

Figure 3: Flow Chart of the System

Many natural language tasks require the accurate assignment of Part-Of-Speech (POS) tags to previously unseen text. Due to the availability of large corpora which have been manually annotated with POS information, many taggers use annotated text to ”learn” either probability distributions or rules and use them to automatically assign POS tags to unseen text.

VI RESULTS AND DISCUSSIONS

Here we have implemented the Sentiment Detection in many ways. For each dataset, the execution of all the methods and in each round and applying different transformations withthe objective to see if have a positive effect on the results. On each round, optimize the method to archive the maximum performance.

Type 1: Results obtained by Reviewing Electronics goods

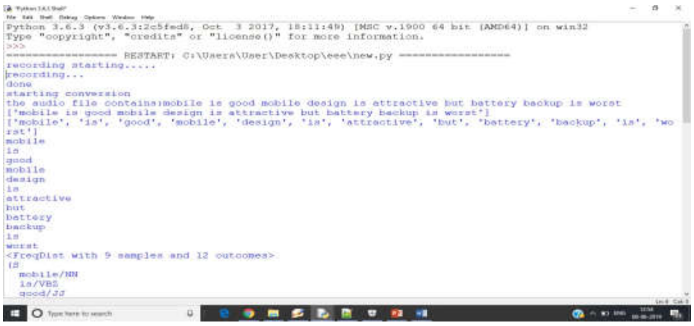

Figure 5: Commenting on Goods

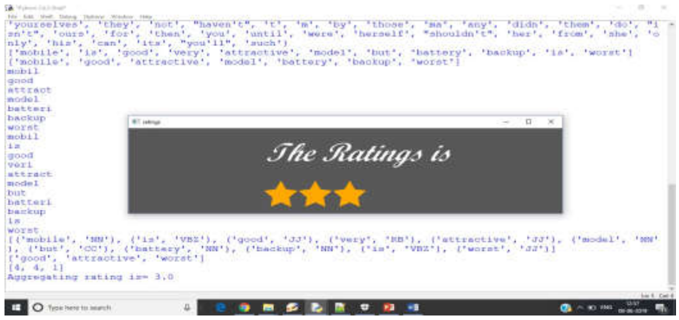

Figure 6:Obtained Reviews

Type-2: Results obtained by Reviewing Fashion goods

Figure 7: Commenting on Dress

Figure 8: Obtained Reviews

Type-3: Results obtained by Reviewing a Movie

Figure 9: Commenting positively on a movie

Positive Comments

Figure 10: Obtained positive Reviews

Figure 11: Commenting negatively on a movie

NEGATIVE COMMENTS

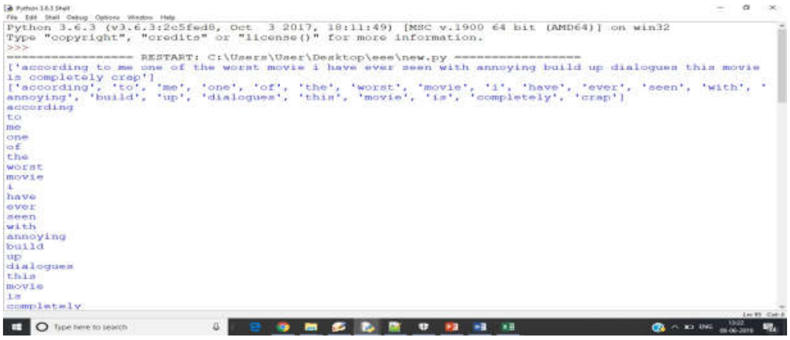

Figure 12: Obtained negative Reviews

VII CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

This system will automatically detects the sentiment in spontaneous natural speech and evaluated this on YouTube, Amazon, Flipkart and many more for audio data. The proposedsystem uses Automatic Sentiment Recognition to obtained transcripts for the audio. A sentiment detection system based on NLTK modeling and POS tagging is used to measurethe sentiment of the transcript. Our various results show it is possible to automatically detect sentiment in natural spontaneous audio with good accuracy. Furthermore, we have also shown that our system is capable of providing key words/phrases that can be used as valuable tags for natural audio.

7.2 Future Scope

Work at character level can make a high impact on the vocabulary size because can be reduced from 50 thousand to, nearly, 3 hundred characters that can represent the most words in the language. Also, is the way that the humans read. Today, the machine learningmethods have problems processing the intermediate meaning behind the time dependencybetween words and, in the same way, the dependency between characters. In Twitter, for example, a sentence has around 20 words that means that the machine needs to learn the dependence between 20 time steps and interpret the hidden sentiment while reading. If wework at character level the machine needs to learn more than 300 hundred time steps. More investigation is needed to make the machines learn like humans reading the sentences at character level.

REFERENCES

[1] A. S. Lakshmish Kaushik and J. H. L. Hansen, "Automatic sentiment detectionin naturalistic audio," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 5, pp. 207–301, December 2016.

[2] Y. X. E. Cambria, N. Howard and T. S. Chua, "Computational intelligence for big social data analysis," IEEE Computational Intelligence Magazine, vol. 11, 2016.

[3] P. S. B. R. S. Cambria, E., "A semantic resource for sentiment analysis based on conceptual primitives," IEEE Intelligent Systems, pp. 2666 2677,, 2016.

[4] L. Barbosa and J. Feng, "Robust sentiment detection on twitter from biased and noisy data," the International Conference on Computational Linguistics, August 2010.

[5] H. X. Z. Zhai, B. Liu and P. Jia., "Clustering product features for opinion mining," Proceedings of Fourth ACM International Conference on Web Search and Data Mining (WSDM-2011), vol. 9, Febraury 2011.

[6] A. S. L. Kaushik and J. H. L. Hansen, "Automatic sentiment extraction from youtube videos," In Automatic Speech Recognition and Understanding (ASRU),IEEEWorkshop, pp. 239–244, 2013.

[7] B. Liu, Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 8 ed., 2012.

[8] B. L. L. Qian, Z. Gao and Y. Zhang, "Automated rule selection for aspect extraction in opinion mining," In International Joint Conference on Artificial Intelligence (IJCAI),2015.