International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

Automated Hybrid Model for Classifying Human Emotions Using Sentiment Analysis and Text Classification

Eze, Chinyere[1], Gabriel, Bariyira Christopher[2], Onuodu, Friday Eleonu[3]

Research Scholar, Department of Computer Science, Ignatius Ajuru University of Education, Rivers State, Nigeria.
Research Scholar, Department of Computer Science, Ignatius Ajuru University of Education, Rivers State, Nigeria.
Visiting Scholar, Department of Computer Science, University of Port-Harcourt, Rivers State Nigeria.

Submitted: 10-07-2021 Revised: 23-07-2021 Accepted: 26-07-2021

ABSTRACT
Emotions are part and parcel of humans. They date as back as the creation itself. Sometimes, emotions are the guiding light to decision making. Therefore, it is imperative that emotions be properly understood and classified to deduce the mood behind certain expressions and actions. Several text classification and sentiment analysis models and techniques have been presented by researchers. However, they had failed to deduce the specific emotions attached to an expression accurately. In this work, we developed an automated hybrid model for classifying human emotions using sentiment analysis and text classification. The system was designed to tackle the inefficiencies associated with the text classification and sentiment analysis approach. The model was implemented using JAVA programming language and MySQlas database. An open source emotion dataset called EMOBANK was used to train our model. We adopted the AGILE methodology in this approach. The results from this system show that texts were accurately classified with an accuracy scoreof 92% and an error of 1.1000%. This outperformed the existing model with an error of 1.211% and accuracy of 79%. This work could be beneficial to psychologists, to companies like AMAZON that operate on digital platforms and to the research communities.

I. INTRODUCTION

Human emotions can be better related to or reacted to appropriately if they were understood. This is vital because this strong feeling defines the motive, mindset and generally the personality of the person who expresses it. Understanding emotions will go a long way in enhancing decision making [1]. In every word, action of gesture of humans, there are emotions embedded in them and most times, these emotions are not clearly visible. Most of human emotions are hidden under sarcasm.

Sentiments aren't always stated in a clear way in content. It is often represented in subtle, complex ways. Besides direct expression of the user's feelings towards a certain topic, he or she can use a diverse range of other techniques to express his or her emotions.Sentiment analysis can also be referred to as opinion mining [2]. The idea behind sentiment analysis is to uncover the opinion or emotions of people towards a particular concept. For example, a company can raise a topic about one of their products online deliberately just to find out what people think about that product. The opinions gathered if properly classified and deciphered can help enhance the decision of the company to improve the quality of the product, make more products or even at worse change the product features. Sentiment analysis also reveals the percentage or degree of the emotions expressed.

Text classification can be used to categorize these opinions, comments and contents to find out what emotions they actually express. The emotions can be happiness, sadness, fear, disgust etc. Text classification

Recently, researcher have grown increasing interest in finding a way to detect subjective information in human texts on social media, blogs, articles, poems, opinions and other contents dropped on these platforms. They have also hoped to detect the emotions behind some words spoken by humans also.

1.1. Statement of the Problem

There are several free text classification and analysis tools which can be used to analyze the

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

emotions from texts and contents on the media and other platforms. However, there is always one limitation or the other with these tools, some of which may include inability to recognize the specific emotions in a text, limited classifications which leaves some vital emotions undetected or wrongly classified under others and lack of efficient models and tools designing the analyzers and classifiers. These limitations will eventually lead to a wrong classification of emotions which in turn can affect decision making.

1.2. Aim and Objectives

The aim of this study is to develop an automated hybrid model for classifying human emotions using sentiment analysis and text classification. The specific objectives are to:

i. develop a text classifier and analysis model for social media comments from a public dataset.
ii. Implement using JAVA programming language and MySQL as database.
iii. evaluate our results with other existing system performance.

II. LITERATURE REVIEW

Text classification and opinion mining also called sentiment analysis are the techniques which form a hybrid tool that was used to develop our system. These concepts will be deeply surveyed in this section for a better understanding of their meaning and applications before the system development proper.

1.3. Emotions: A Survey

Feelings seem more primitive than even thoughts. The saying "I feel, therefore I am" moves to promote this claim. Emotions were at the very inception of life itself and has remained part and parcel of humans ever since. Emotions have been defined in several ways. We shall explore some of these definitions. The first definition is by Garret [3] she defined emotions as feelings, experience, physiology, behavior, cognitions and conceptualizations. Wikipedia defined emotions as a mental and physiological state associated with a variety of feelings, thoughts and behavior.Gaind et al [4] defined emotion as intense feelings that are directed at something or someone in response to internal or external events having a particular significance for the individual.From all these definitions, feelings are a repeatedly component of the definitions. This therefore explains that emotions are merely the feeling behind an action or word expressed by someone about someone else or a concept.Emotion can be expressed in many ways that can be seen such as facial expression and gestures, speech and by written text [5].

Emotions form a very important and basic aspect of our lives. Emotions are reflected in what we do i.e. our actions, what we say our words and expressions though may not be directly [6]. Our emotions also affect our daily lives because most emotions are driven by moods. Emotions are complex which makes them difficult to decipher or understand most times. An emotional expression towards a person or an object will spark a kind of reaction from that person or controller of the object this is why it is important to always understand the emotions behind a word or action before reacting to it to avoid an inappropriate reaction.

The social media has been crowned the chief place for emotion deposit and withdrawal. A lot of emotions are poured out daily on these platforms towards topical issues of politics, relationships, education, businesses, values etc. Twitter, Facebook, Telegram, YouTube and Blogs are some of the platforms which create a forum for comments and expression of views concerning any issue dropped in the platform. University portals also create this platform for reviews and suggestions about the activities in the school, reports of misconducts and other complaints. All these are sources of opinions which carry emotions that need to be analyzed and addressed.

1.3.1. Emotional Models

There are some emotional models in existence. They are briefly described in this section.

The Ekman's basic emotion model andthe Plutchik's bipolar emotion model fall in this category.Ekman in his model has divided emotions into six discreteclasses of anger, disgust, fear, happiness, sadness and surprise. Plutchik's model, on the otherhand, is a super set of Ekman's model with two additionalclasses: TRUST and ANTICIPATION.The emotional dimensions follow the approach ofrepresenting the emotion classes in a dimensional form: either2D or 3D, with each emotion occupying a distinct position inspace. These emotions can be described in 3 dimensions of:pleasurable or unpleasurable, arousing or subduing and strainor relaxation. The three dimensions could also be defined as:pleasant vs. unpleasant, attention vs. rejection and level ofactivation. Some research has been done using both themodels of emotion representation.

Most dimensional models have valence and arousal orintensity dimensions: valence dimension indicates how muchpleasant or unpleasant an emotion is, arousal dimensiondifferentiates between activation and deactivation states. Themost prominent ones being as defined under:

2D models:

i. **Circumplex model:** defined this model where thevertical axis represent the arousal and horizontal axisrepresent the valence, while the origin representsneutral valence and a medium level of arousal.

ii. **Vector model:** consist of two vectors

Pointing intwo directions assuming the presence of an underlyingarousal dimension with valence dimension vectordetermining the direction in which a particular emotionlies.

iii. **PANA(positive activation-negative activation) model:** it divides the

system into positive and negativeeffect, with vertical axis representing the low to highpositive affect and the horizontal axis representing thelow to high negative effect.

3D models

i. **Plutchik' model:**Plutchik gave a hybrid

modelarranging the emotions into concentric circles withinner being the basic and the outer more complexemotions.

ii. **PAD (pleasure, arousal and dominance) model:** In addition to

arousal and valence, it describes a thirddimension of the dominance, which indicates whetherthe subject feels in control of the situation or not.

iii. **Lovheim cube model:** presents

Signalsubstances forming the axis of the coordination systemand the eight basic emotions are placed at the eightcorners of the cube.

1.4. Emotion Detection by Text Classification

Emotion Detection in text documents is essentially a content based classification problem involving concepts from the domains of Natural Language Processing as well as Machine Learning [5].

Emotion detection (ED) is a branch of sentiment analysis that deals with the extraction and analysis of emotions [7]. Emotion detection is basically aimed at uncovering the underlying emotions behind the writers' or writer's words by examining the input texts. This is based on the assumption that if a person's emotions are good or happy, the choice of words might comprise of positive texts. Likewise, if a person is sad, frustrated or angry, the kind of words they use can allow us to infer their underlying negative emotion [1].

Emotion detection in text has a number of important applications. Some of them include: in the area of business development, emotion detection can influencedevelopment of strategies by marketers for customer relationship and vice versa. Psychologists can benefit from being able to infer people's emotions based on the text that they write which they can use to predict their state of mind. This knowledge can be practically applied to predict consumer behaviour and customer preferences for corporate financial gain. In the field of education, the ability of computers to automatically track attitudes and feelings with a degree of human intuition has contributed to the development of Text-to-Speech systems and Intelligent Tutoring Systems (ITS).

But there are some complicated emotions which cannot be classified as either positive or negative just by examining the input words.

1.5. Sentiment Analysis

Sentiment analysis is a type of natural language processing for tracking the mood of the public about a particular product or topic. Sentiment analysis, which is also called opinion mining, involves in building a system to collect and examine opinions about the product made in blog posts, comments, reviews or tweets. Sentiment analysis can be useful in several ways. For example, in marketing it helps in judging the success of an ad campaign or new product launch, determine which versions of a product or service are popular and even identify which demographics like or dislike particular features [8].

The emotion analysis uses the natural language processing, text analysis and various computational techniques to determine the emotion

s hidden in a particular text. This analysis can be done at various levels: document level [9] sentence level,word level, and aspect level. The emotion analysis of some input data consists of the following steps as shown in figure 1.

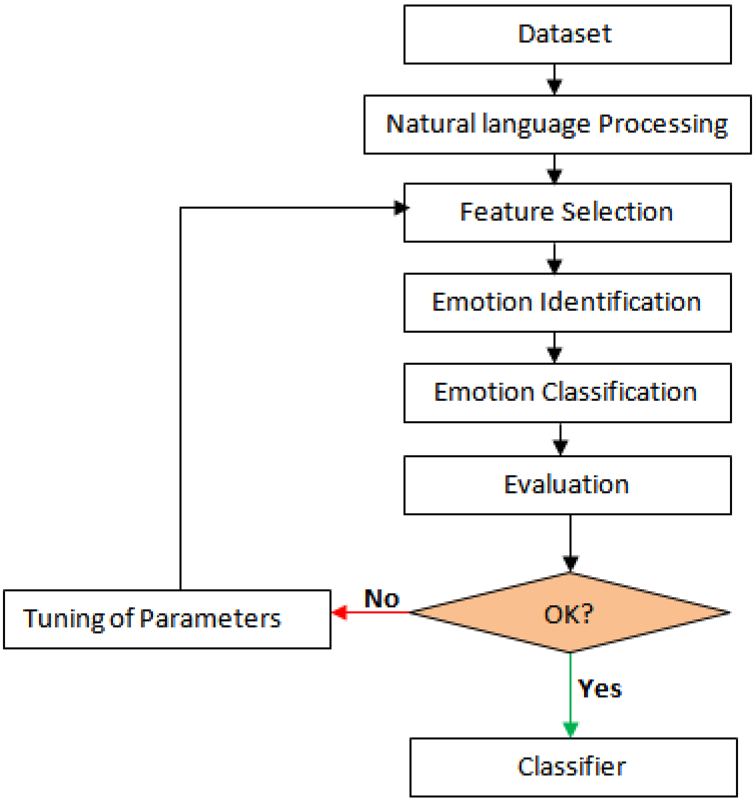

Fig. 1: Steps for Emotion Analysis (Source: Hakak et al [6])

The field of sentiment analysis is one that is rapidly growing and lies between linguistics and computer science. Sentiment analysis is used to broadly classify a language as positive or negative [10]. Sentiment analysis is regarded as a crucial task for several application domains, including business, social well-being, politics, security, and software engineering. Emotion analysis form a fundamental aspect of affective computing. There are several tools which are used for sentiment analysis. The TWEEPY and VADER tool for scooping and analyzing twitter texts to mine the emotions embedded in the opinions tweeted by the users are examples of the sentiment analysis tools.

1.6. Techniques for Emotion Classification

There are basically four techniques for emotion detectionincluding Keyword spotting technique, Lexical affinitymethod, Learning based method and Hybrid method.

i. Keyword Spotting

This technique can be defined on the basis of some definitepredefined keywords which have been classified into somecategories such as surprised, angry, sad, joy and disgusted.This technique analyzes the intensity of words to identifyemotion.

ii. Lexical Affinity

Lexical Affinity approach works with probabilistic 'affinity'for a specific opinion to random words spaced out frompicking up sentimental keywords.

iii. **Learning and Hybrid method**

Learning based method is used to categorize the input stringinto different emotions. Learning based methodology andkeyword based spotting methodology is combined to createHybrid methods.

iv. Hybrid methods:

Since keyword-based methods with thesaurus and naïve learning-based methods could not acquire satisfactory results, some systems use hybrid approach by combining both keyword spotting technique and learning based method, which help to improve accuracy. The most significant hybrid system so far is the work of Wu, Chuang and Lin [11], that utilizes a rule-based approach to extract semantics related to specific emotions and Chinese lexicon ontology to extract attributes. These semantics and attributes are associated with emotions in the form of emotion association rules. As a result, these emotion association rules, replacing original emotion keywords, serve as the training features of their learning module based on separable mixture models. This method outperforms previous approaches, but categories of emotions are still limited.

1.7. Related Work

Binali et al [1] proposed computational approaches for emotion detection in text. They demonstrated how these modelscould be used by discussing computational approaches to emotion detection. They proposed a hybrid based architecture for emotion detection. The SVM algorithm was used for validating the proposed architecture and achieved a prediction accuracy of 96.43% on web blog data.

Calefato et al [12] proposed a toolkit for emotion recognition from text. They presentedEmoTxt, an open source toolkit for emotion detection from text, trained and tested on two large gold standard datasets mined from Stack Overflow and Jira. They released the classification models to be used for emotion detection tasks. EmoTxt also supported training of emotion classifiers from manually annotated training data. From text, trained and tested on two large gold standard datasets mined from Stack Overflow and Jira. Other than classification, EmoTxt supports training of emotion classifiers from manually annotated training data.

Gaind et al [4] presented emotion detection and analysis on social media. They proposed a method for classifying input text into six categories of emotions such as happiness, fear, disgust, surprise, sadness and anger. They combined two approaches to ease the extraction of emotions from text. The first approach was natural language processing and the second approach was machine learning. They were able to develop bags of emotion al words along with their emotional intensities. Their model showed some level of accuracy in classifying selected twitter words.

Erik et al [13] presented automatic sentiment analysis in On-line text. They provided an overview of the various techniques used to tackle the sentiment analysis crisis. They also indicated the usefulness of the sentiment analysis as a model for text (opinion) extraction. However, their research failed to cover approaches that can be used to mine opinions from several sources on the World Wide Web such as blogs etc.

Bhowmick et al [14] presented a study on classifying emotions in news sentences into emotion categories. They conducted different experiments using 1000 corpus to compare human classifications to machine classification into categories such as happiness, sadness, fear, anger, disgust and surprise. Their result showed that best classification was achieved when anger and disgust classes were combined and surprise class was removed. The average precision was computed to be 79.5% and the average class wise micro F1 was found to be 59.52%.

Shivhare and Khethawat [5] presented a research on emotion detection from text. They discussed emotion recognition based on textual data and techniques used in emotion detection. They also discussed the limitations of these techniques in their application to emotion detection. However, they could not develop a model to enhance emotion detection and solve the existing problems.

Mohammad [15] proposed a study on sentiment analysis for detection valence emotions and other affectual state from texts. They outlined diverse landscape of applications and problems associated with automatic sentiment analysis. Key features, algorithms and datasets used in sentiment analysis were also outlined. They also discussed the initial approaches for handling sentiment modification by Negators and modals.

Acheampong et al [7]proposed text-based emotion detection. They carried out a survey on existing literatures on emotion detection. They related literatures were surveyed in relation to their contribution, dataset used, approaches/techniques employed and their results attained at the end of the day. Their strengths and weaknesses were also discussed. Finally, they presented emotion labeled data sources to provide neophytes with eligible text dataset for emotion detection.

Hasan et al [2] proposed a machine learning based sentiment analysis for twitter accounts. They provided a comparative analysis of techniques used for sentiment analysis for political view analysis using supervised machine learning algorithms such as Naïve Bayes and support vector machines (SVM). However they were unable to implement a model for enhanced sentiment analysis.

Molina-Gonzal et al [16]presented a study to analyzethe emotions of textual dialogue from SemEval-2019 Task3: EmoContext for English language by various systems. Their main contribution is the integration of emotional and sentimental features in the classification using the SVM algorithm. However, they did not develop any model for this classification.

Gottipati et al [17]presented an automated solution for extracting the explicit suggestions from the students' qualitative feedback comments. The implemented solution leveraged on existing text mining and data visualization techniques. It comprised of three stages, namely data pre-processing, explicit suggestions extraction and visualization. Their solution was evaluated using student feedback comments from seven undergraduate core courses taught at the School of Information Systems, Singapore Management

University. They comparedrule-based methods and statistical classifiers for extracting and summarizing theexplicit suggestions. Based on our experiments, the decision tree (C5.0) worked thebest for extracting the suggestions from students' qualitative feedback.

Nahar et al [18] presented a review of sentiment analysis and emotion extraction. They represented the study of sentiment analysis to extract emotions from textual contents by using several methods of tracking mode. Several classifiers were trained to extract basic moods such as Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger, and Anticipate from text. The aim of their research was to give a nearly concept about sentiment analysis techniques and the related areas.

Hutto and Gilbert [19] proposed VADER: a parsimonious rule-based model for sentiment analysis of social media text. Using a combination of qualitative and quantitative methods, they constructed and empirically validated a gold-standard list of lexical features (along with their associated sentiment intensity measures) which were specifically attuned to sentiment in micro blog-like contexts. Next, they combined these lexical features with consideration for five general rules that embody grammatical and syntactical conventions for expressing and emphasizing sentiment intensity. They observed from their results that using the parsimonious rule-based model to assess the sentiment of tweets, VADER outperforms individual human raters (F1 ClassificationAccuracy = 0.96 and 0.84, respectively), and generalized more favorably across contexts than any of our benchmarks.

Xiao et. al.[20] presented a study to have extract and analyze a featureset by "Fisher's Discrimination Ratio" and selected few bestfeatures that are computed from F0-fundamental frequencywith range between 60-450 Hz. Including the feature list andselection result. They introduced a new definition ofemotions as 3-states emotion with active, passive and neutral state. They also built a database to verify states ofemotional speech with speech samples.

III. METHODOLOGY

Agile methodology was adopted in this approach the system. The agile methodology is an iterative and incremental based development where requirements are changeable depending on the needs of the customer. Agile method helps in adaptive planning, time boxing and development in an iterative manner. This methodology is a theoretical framework that promotes foreseen interactions throughout the development cycle. The Agile method follows the same processes that are outlined in the SLDC such as the requirements gathering, analysis, design, coding testing and delivers a prototype of the design to the client while waiting for his/her response before delivering the complete software which will contain the user's modified corrections and requirements.

1.8. Analysis of the Existing System

The existing system was proposed by Gaind et al [4]. They hoped to address the problem of detection, classification and quantification of emotions of texts in any form. They considered texts scraped from social media platforms like Twitter. They proposed a model which classified text into 6 emotion categories (Happiness, Sadness, Fear, Anger, Surprise and Disgust). They combined natural language processing (NLP) and machine learning classification algorithms in this approach. Their models supported automatic creation of training set thereby elimination manual annotations of large datasets. They successfully created a large bag of emotional words along with their emotion intensities. With their model, they could score and label any piece of texts especially texts from twitter posts. They also developed bags of English emotional words and attached intensities to those words. They adopted an open source classifier called J48, and an open source NLP tool called CoreNLPand with these tools carried out text classification of large datasets. Their model was also used to classify emojis included in the text. The intensities used to quantify the emotions were, strong, light and medium. They introduced the concept of Surety Factor to suggest the reliability of their results and the degree of usefulness and correctness of the results. The results were visualized using pie-charts, bar graphs and maps.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

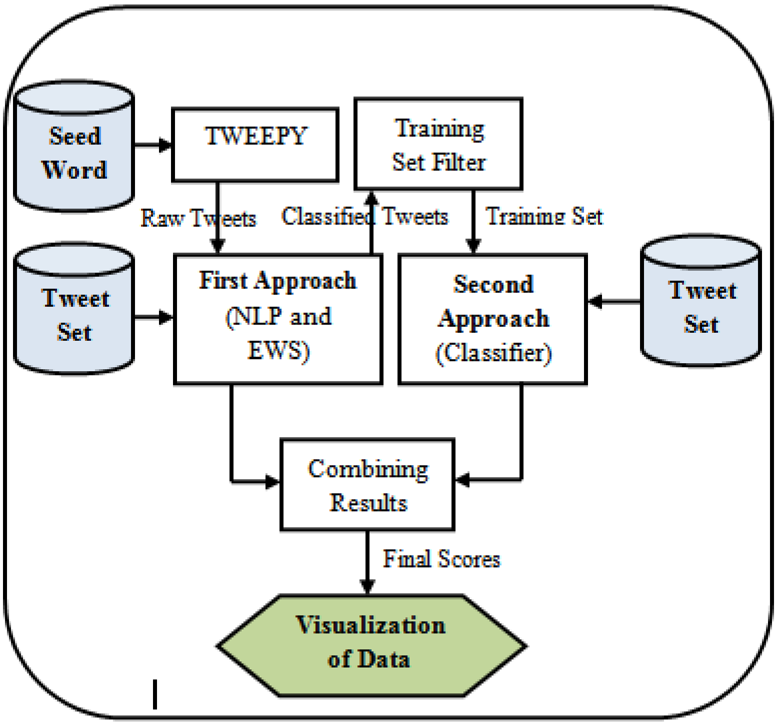

Fig. 2: Architecture of the Existing System (Source: Gaind et al [4])

1.8.1. Disadvantages of the Existing System

The existing system however has the following drawbacks.

i. The existing model made use of already designed tools for text classification instead of a well-defined tool created using their system specification. Therefore their results cannot be termed efficient as they claim.

ii. The existing model could not automatically update the bag of words which was previously created.

iii. The proposed model could not cover otherforms of hidden emotions which cannot be clearly detected by mere text analysis.

1.8.2. Algorithm of the Existing System

Step 1:

Declare Variables;

TS	=	Tweet Sets
EW	=	Emotion words
TS	=	Tweet sets
TSF	=	Training Set Filter

Step 2:

Scrape Tweets using TWEEPY;

Create TS;

Step 3:

Apply coreNLP+ EW on TS;

Step 4:

Launch Training Set Filter;

Step 5:

Launch J48 classifiers;

Step 6:

Combine Results;

Results = CoreNLP + EW + J48;

Step 7:

Display Results in Charts, Graphs & Maps

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

1.9. Analysis of the Proposed System

The proposed model is an improvement of the existing system. The problem solved by the proposed system can be represented in this expression:

IF S: Group of words (excellent, poor, improve)

And N: Sentiment categories (negative, positive and neutral).

THEN $S = 2^n$?

The model is aimed at classifying emotions accurately with minimum error. Our model made use of text gotten from an online dataset called EMOBANK. This dataset contains texts from various sources such as blogs, fictions etc. The emotions contained in the text include happiness, sadness, Joy, fear, anticipation, love, anger and disgust. This is the largest emotion category ever defined by any sentiment analyzer. These emotions contain other emotions embedded in them as follows: Disgust (boredom, loathing), anger (annoyance, rage and aggressiveness), anticipation (vigilance, interest and optimism), Love (admiration, trust and acceptance), Surprise (awe, amazement and distraction), Fear (terror, apprehension, submission) and sadness (disapproval, pensiveness and grief). Other emotions such as remorse lie between disgust and sadness. First our system defines rules where a collection of words are used to categorize texts as positive/good, negative/bad and average/neutral. Note that a text can be characters, words, phrases, sentences and emojis. NLP is applied to this text. Specifically Naïve Bayes algorithm which promotes independence of set of features being classified. The algorithm is defined as:

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$

$$\text{i.e., Posterior} = \frac{\text{Prior} * \text{likelihood}}{\text{Evidence}}$$

This means that any given text can either be positive or negative or neutral, but can never be both or three of them at the same time. This algorithm is more effective in this approach. This algorithm is used to train a classifier with separate class labeled data. Table 1 shows an example of this.

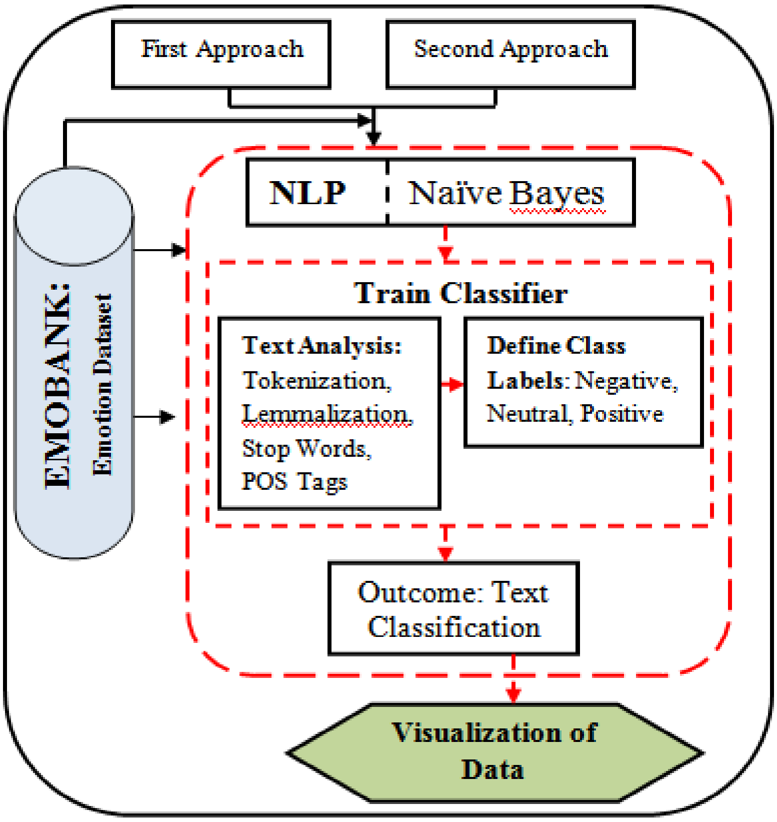

Fig. 3: Architecture of the Proposed System.

Table 1: Class Label Classification

REVIEW	CLASS LABEL
This is a very good product	Positive/Good
Good product. Taste could be better though.	Average/Neutral
Ewww!!! I regret buying this product. Whack!!!	Negative/Bad

The next step involves text analysis, involving tokenization, POS tagging, removal of stop words etc. This is the data that will be used for the training of a classifier, so that it can automatically detect the categories of text in a large pool of data and classify them accurately. During the training polarity values are assigned to the input words ranging from -1 to 1. -1 is assigned to negative words, 0 is assigned to neutral words and 1 is assigned to positive words.

Therefore, when the system is given an input word, it carries out sentiment analysis to mine the emotions embedded in the text tags them in their groups and moves to classify the text according to the defined rules. The results from these two processes are merged to give the rating of that text based on the classification.

1.9.1. Advantages of the Proposed System

The proposed system has the following advantages:

i. The proposed model developed a model using self-defined specifications for efficient results.
ii. The proposed model combined two techniques to effectively classify text based on emotions classes and text categories.
iii. The proposed model covered other forms of hidden emotions which cannot be clearly detected by mere text analysis.

1.9.2. Algorithm of the Proposed System

Step 1:
Start;
Step 2:
Declare Variables;
EW = Emotion words
TSF = Training Set Filter
ET = Emotion Tags
NBA = Naïve Bayes Algorithm
ED = Emotion Dataset
SW = Stop Words;
CL = Class Labels
Step 3:
Var ED As Input;
ED = EW + ET
Step 4:
Apply NBA;
$$NBA \longrightarrow P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$
Step 5:
Launch Training Set Filter;
Step 6:
Train Classifiers;
Tokenize EW;
Remove SW;
Step 7:
Define CL;
CL= (For EW=0; EW<=1; EW++)
Step 8:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

Outcome: Negative (-1)/Positive (1) & Neutral (0)
Results = CoreNLP + EW + J48;
Step 9:
Display Results in Charts, Graphs & Maps
Step 10:
Quit System;

IV. IMPLEMENTATION

The system was implemented using JAVA programming language. Java is an object oriented programming language that supports machine learning programming and training of large datasets for supervised and unsupervised learning models. Netbeans was the IDE used in developing the program for this implementation. The model adopted a dataset from a public repository of emotion datasets called EMOBANK. This dataset consists of over 10,000 sentences annotated dimensionally in accordance to the valence arousal dominance (VAD) emotion representation model. The sentences are sourced from blogs, newspapers, letters, travel guides, fiction, news headlines etc. which spans a wider domain. The emotions annotated in this dataset include anger, happiness, joy, anticipation, disgust, surprise, hate and love. This dataset was used as the training set together with the defined text classification categories.
A sample of the dataset is illustrated in table 2:

Table 2: Dataset Express from ECOBANK

SENTENCES	SENTIMENT TAG	INTENSITY
The country is under fire since this administration.):	Fear[1], Anger[2], Disgust[3]	Mild, Mild, Mild
Poetry Saved my Life!!	Joy[1], Happiness[2], Love[3]	Strong, Mild, Mild
Trump is the worst	Sadness[1], Hate[2]	Strong, Strong
I look forward to this years' summer vacation.	Anticipation	Very Strong.
I don't know how I feel about this..	Sadness	Mild
Hahahahahahaha!!!	Happiness,	Strong
Women rule this world	Hate	Rage

1.10. Discussion of Results

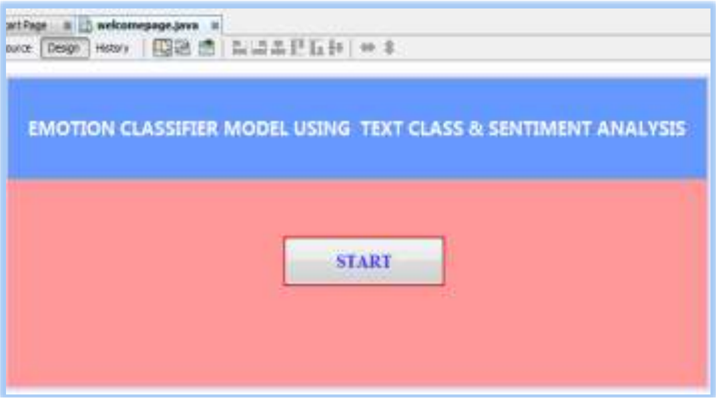

Figure 4: Welcome Page: Emotion Classifier

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

Figure 4 displays the Netbeans IDE used for the development of our model. It also displays the welcome page. Certain Java classes were imported to support chart creation to display our results in a visualized form. Some keywords were defined and assigned intensities from 1-10, if the intensity value is high i.e. greater 5, it implies that the emotion is negative and strong words such as terrible, disgusting, worst, useless were included in the input text. However, if the intensity is less than 5, it implies that the emotions or opinions are positive and good words such as amazing, beautiful, awesome etc were used in the input sentence.

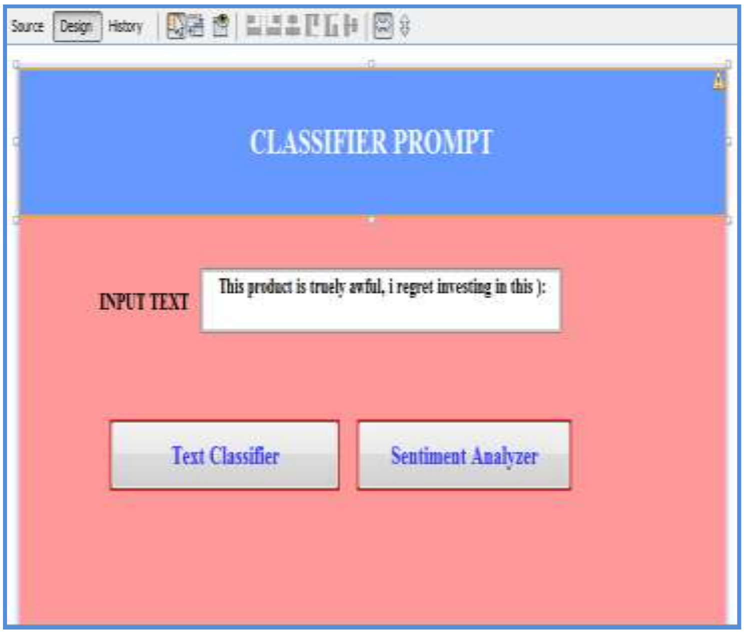

Figure 5: Input for Text Classification: Emotion Classifier

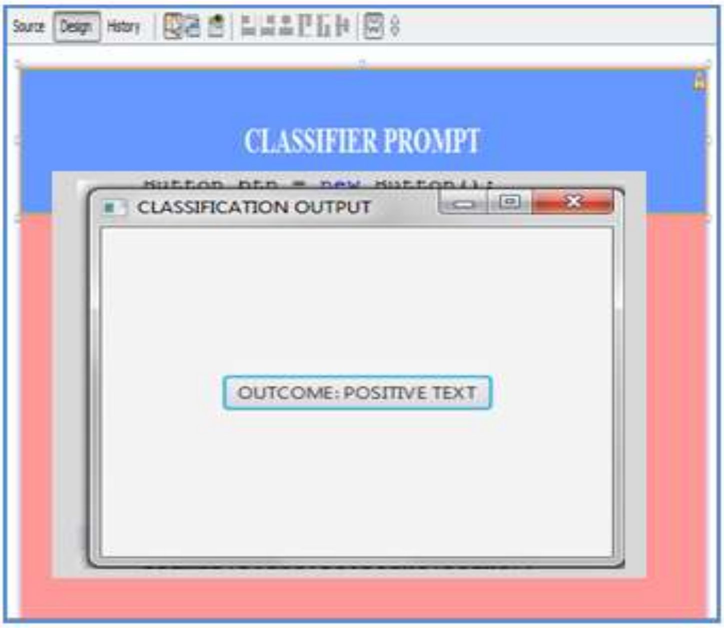

Figure 6: Text Classification Output: Emotion Classifier

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

Once a text is inputted in the textbox (this could be a phrase, word, paragraph, word, character or emoji), and the classify button is clicked, the model automatically and immediately provides a prompt to display the class label of the text, i.e. negative comment, positive comment or neutral comment.This works on the basis of the Naïve Bayes algorithm, and the classification is independent of all other classifications. Then this text is stored in the database along with its class label.

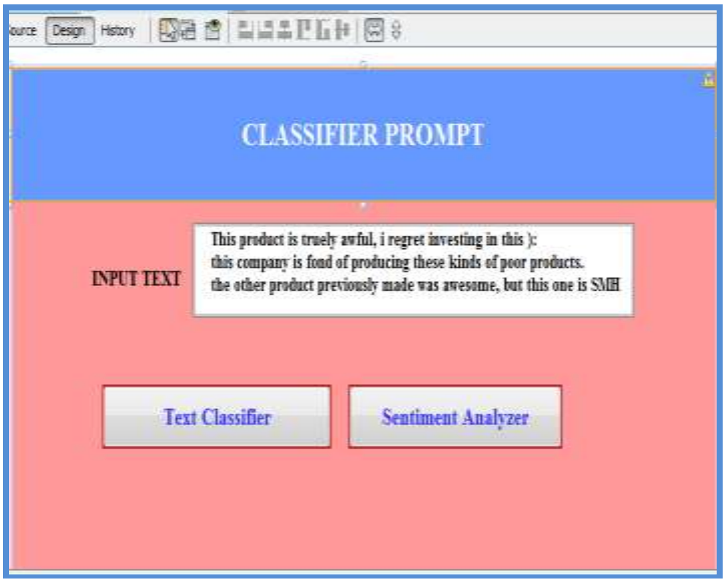

Figure 7: Input for Sentiment Analysis

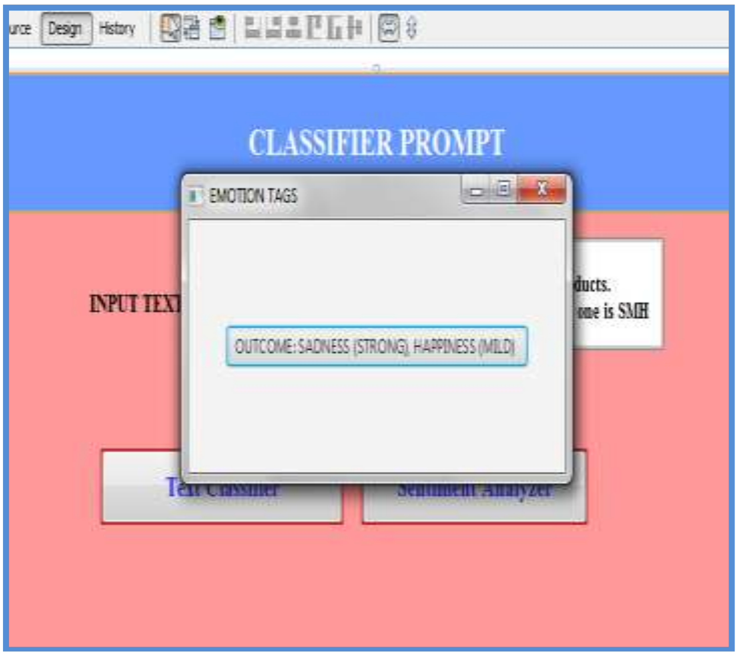

Figure 8: Sentiment Analysis Output

For sentiment analysis, once a text is typed in, and sentiment analysis is clicked, the model refers to the EMOBANK and scans for emotion tags on similar texts stored in the BANK, and then it assigns the best match emotion tag to the text. Several tests were carried out and the outcomes were accurate.

Finally, a pie chart is constructed to show the distribution of various emotions present in the text,

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2956-2971 www.ijaem.net ISSN: 2395-5252

each represented by a slice according to its intensity. This is shown in figure 9.

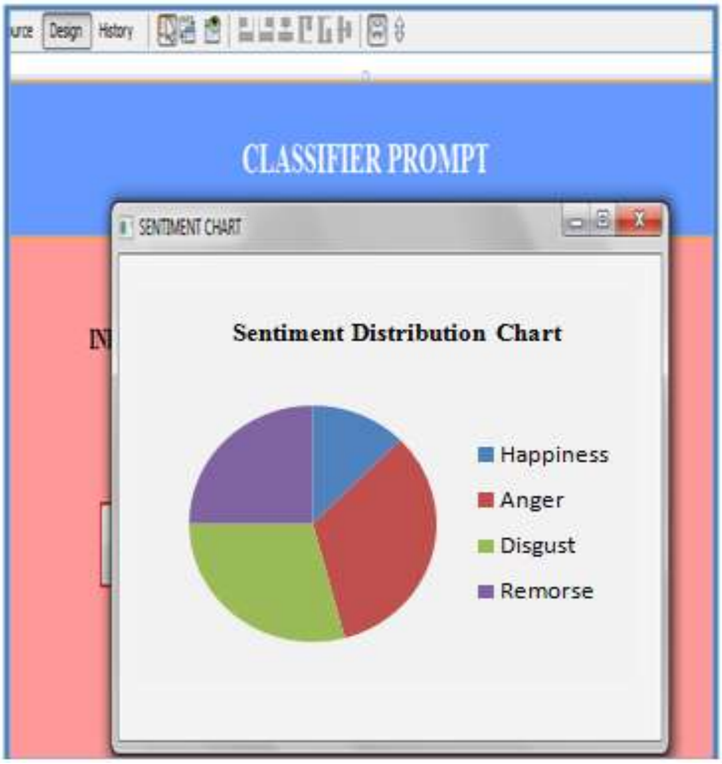

Figure 9: Sentiment Distribution Chart

1.11. Performance Evaluation

After the system implementation, testing was carried out to evaluate the performance of the model in text classification.Certain parameters were used to carry out this evaluation such as how fast the classification occurred, how efficient the classification was based on the testing set. We used about 10,000 training datasets to train our model and defined a set of emotion word and classification to train our classifier. We used about 22 data to test our system and the results were exactly accurate. Less classification error was recorded as opposed to that of the existing system with larger error rate.

Using Gainds et al [4] error rate calculation based on the parameters defined in table 3.

Error rate= Error rate= $\frac{\text{No. Test Data}}{\text{No. of Correct Classifications}}$

$= \frac{22}{21} = 1.100\%$

Table 3: Performance Evaluation

S/N	Parameter	Existing Model	Proposed Model	Parameter
1	No. of Training Data (NTrD)	3,200	10,000	No. of Training Data (NTrD)
2	No. of Testing Data (NTsD)	20	22	Mean Square Error (MSE) (NTsD)
3	Classification Speed (CP)	10 Minutes	1 Minute	Classification Speed (CP)

4	Classification Accuracy (CA)	10%	25%	Classification Accuracy (CA)
5	Cross Platform Adaptability (CPA)	10%	20%	Cross Platform Adaptability (CPA)
6	Model Efficiency (ME)	15%	20%	Model Efficiency (ME)
7	Dataset Quality (DQ)	15%	18%	Dataset Quality (DQ)
8	Classification Error (CE)	1.210%	1.100%	Classification Error (CE)

Total PerformanceScore = $\frac{\sum \text{Parameters}}{\text{Error Rate}}$

Existing System = $\frac{5+20+10+10+10+15+15}{1.211}$

= 79.77%

Proposed System = $\frac{10+22+1+22+20+24.5+18}{1.1}$

= 97.56.%

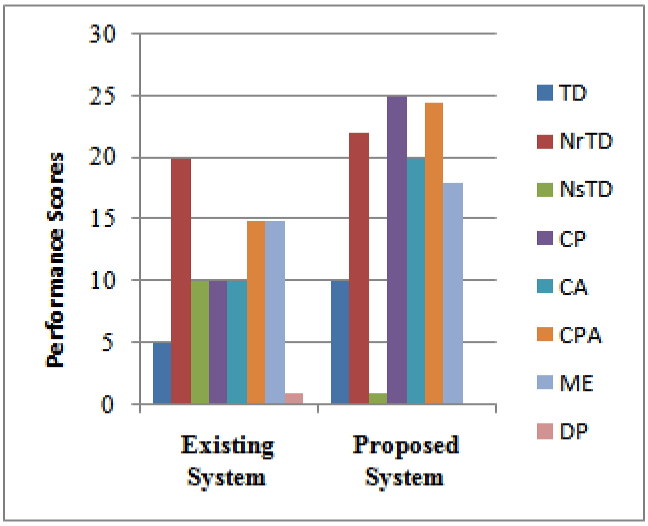

Fig. 10: Performance Evaluation Chart.

V. CONCLUSION

We have successfully developed a model to accurately and automatically classify input words according to text classification categories such as negative, positive and neutral. Text classification is a very useful approach in various sectors. It helps companies and organization lay hands on the opinions of their customers on a large scale. Sentiment analysis also called opinion mining is used to mine these opinions and filter the emotion or mode of the audience when expressing these opinions. An open source emotion bank with 9 annotated emotions such as hate, anger, joy, happiness, disgust, anticipation, love, sadness and surprise was adopted for data training and testing in this model. Our results show that our system is more accurate than the other existing systems.

CONTRIBUTION TO KNOWLEDGE

An automated hybrid model for classifying human emotions using sentiment analysis and text classification has been developed. This model is more efficient than other existing text classification and sentiment analysis algorithms.

SUGGESTION FOR FUTURE WORK

For future work, we suggest thatan automated model that auto-updates emotion words bags and tests with a larger datasetof above 10,000 texts. Be developed.

REFERENCES

[1] H. Binali, C. Wu and V. Potdar."Computational Approach for Emotion Detection in Text." 4th IEEE International Conferences on Digital Ecosystem and Technologies (IEEE DEST 2010). Pp. 172-177. MAY 2010

[2] A. Hasan, S. Moin, A. Karim and S. Shamshirband. "Machine Learning-based Sentiment Analysis for Twitter Accounts." Mathematical and Computational Applications. Vol. 23, No. 11, pp. 1-15. Doi: 10.3390/mca23010011. FEB. 2018.

[3] B. Garett, "Brain and Behavior: an introduction to biopsychology." 2nd ed.: SAGE Publications, 2009.

[4] B. Gaind, V. Syal and S. Padgalwar. "Emotion Detection and Analysis on Social Media." Journal of Engineering Science and Researches. Pp. 78-89, MAY 2018. http://www.gjesr.com/ICRTCET-18.html)

[5] S. N. Shivare and S. Khethawat. "Emotion Detection from Text."Research gate Publication. Pp. 1-7. 2017.

[6] N. M. Hakak, M. Modh, m. Kirmani and M. Mohd. "Emotion Analysis; A Survey."2017International Conference on Computer, Communicationsand Electronics (Comptelix) Manipal University Jaipur, Malaviya National Institute of Technology Jaipur & IRISWORLD. Pp. 397-402.

[7] F. A. Acheampong, C. Wenyu and H. N. Mensah. "Text-Based Emotion Detection: Advances, Challenges and Opportunities." Wiley. Pp. 1-21. APR. 2020. Doi: 10.1002/eng2.12189

[8] G. Vinodhini and R. M. Chandrasekaran. "Sentiment Analysis and Opinion Mining: A Survey." International Journal of Advanced Research in Computer Science and Software Engineering. Vol. 2, Issue 6, pp. 282-292. JUN. 2012

[9] R. Moraes, J. V. Francisco and W. P. Gavião "Neto-Document-level Sentiment classification: an empirical comparison between SVM and ANN." Expert Syst Appl, pp. 621–633. 2013.

[10] M. Taboada. "Sentiment Analysis: An Overview from Linguistics." Annual Review of Applied Linguistics. Pp. 1-45. 2016

[11] C.-H. Wu, Z.-J. Chuang and Y.-C. Lin, "Emotion Recognition from Text Using Semantic Labels and Separable Mixture Models," ACM Transactions on Asian Language Information Processing (TALIP), vol. 5, issue 2, Jun. 2006, pp. 165-183, doi:10.1145/1165255.1165259.

[12] F. Calefato, F. Lanubile and N Novielli. "EmoTxt: A Toolkit for Emotion Recognition from Text." Pp. 1-2.

[13] B. Erik, H. Pieter, K. Deschacht and M. F. Moens. "Automatic Sentiment Analysis in On-line Text." Proceedings ELPUB2007 Conference on Electronic Publishing –Vienna Austria. Pp. 349-358. JUN 2007

[14] P. K. Bhowmick, A. Basu and P. Mitra. "Classifying Emotions in News Sentences: When Machine Classification meets Human Clasification." International Journal on Computer Science and Engineering.Vol.2 Issue 1, pp. 98-108. 2010.

[15] S. M. Mohammad. "Sentiment Analysis: Detecting Valence, Emotions and other Affectual States from Text."Emotion Measurement 2015. Pp. 1-40.

[16] F. M. Plaza-del-Arco, M. D. Molina-Gonzal, M. T. Martin-Valdivia and L. A. Urena-Lopez. "SINAI at SemEval Task 3: Using Affective Features for Emotion Classification in Textual Conversation" Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019), Minneapolis, Minnesota, USA, pp. 307–311. JUN. 6–7, 2019

[17] S. Gottipati, V. Shankararaman and J. R. Lin. "Text Analysis Approach to Extract Course Improvement Suggestions from Student's

Feedback." Research and Practice in Technology Enhanced Learning. Vol. 13, No. 6. Pp. 1-19. 2018
[18] L. Nahar , Z. Sultana, F. Tasnim and A. Chowdhury. "Sentiment Analysis and Emotion Extraction: A Review of Research Paradigm." Research gate Publication. pp. 2-8
[19] C. J. Hutto and E. Gilbert. "VADER: A Parsimonious Rule-Based Model for Sentiment analysis of Social Media Text." Association for the Advancement of Artificial Intelligence. Pp. 1-11. 2014
[20] Z. Xiao, E. Dellandrea, W. Dou, and L. Chen, "Features extraction and selection for emotional speech classification." In Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on, 2005, pp. 411-416.