# COMPUTATIONAL RECOGNITION OF EMOTIONS: SENTIMENT ANALYSIS IN HUMAN COMPUTER INTERACTION

**Thesis Submitted**
**in Partial Fulfillment of the Requirements for the Degree of**

## MASTER OF TECHNOLOGY

**in**

**INFORMATION TECHNOLOGY**

**by**


**NAINA SHARMA**
**2K23/ITY/06**


**Under the supervision of**


**PROF. DINESH KUMAR VISHWAKARMA**
Professor, Department of Information Technology
Delhi Technological University



**DEPARTMENT OF INFORMATION TECHNOLOGY**
**DELHI TECHNOLOGICAL UNIVERSITY**
**(Formerly Delhi College of Engineering)**
**Shahabad Daulatpur, Main Bawana Road, Delhi 110042, India**


**May, 2025**

**DEPARTMENT OF INFORMATION TECHNOLOGY**
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

# ACKNOWLEDGEMENT

We wish to express our sincerest gratitude to Prof. Dinesh Kumar Vishwakarma for his continuous guidance and mentorship that he provided us during the project. He showed us the path to achieve our targets by explaining all the tasks to be done and explained to us the importance of this project as well as its industrial relevance. He was always ready to help us and clear our doubts regarding any hurdles in this project. Without his constant support and motivation, this project would not have been successful.

Place: Delhi — Naina Sharma

Date: 30.05.2025 — (2K23/ITY/06)

**DEPARTMENT OF INFORMATION TECHNOLOGY**
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

## CANDIDATE’S DECLARATION

I, Naina Sharma, 2K23/ITY/06 students of M.Tech (Information Technology), hereby certify that the work which is being presented in the thesis entitled “Computational Recognition of Emotions: Sentiment Analysis in Human Computer Interaction” in partial fulfilment of the requirements for the award of degree of Master of Technology, submitted in the Department of Information Technology, Delhi Technological University is an authentic record of my own work carried out during the period from Jan 2025 to May 2025 under the supervision of Prof. Dinesh Kumar Vishwakarma.
The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.

**Candidate’s Signature**

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

**Signature of Supervisor (s)**

**DEPARTMENT OF INFORMATION TECHNOLOGY**
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

# CERTIFICATE

I hereby certify that the Project Dissertation titled “Computational Recognition of Emotions: Sentiment Analysis in Human Computer Interaction” which is submitted by Naina Sharma, Roll No – 2K23/ITY/06, Department of Information Technology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Prof. Dinesh Kumar Vishwakarma
Professor

Date: 30.05.2025

Department of Information Technology
Delhi Technological University

# ABSTRACT

In today’s digitally-driven world, Human-Computer Interaction (HCI) has evolved far beyond simple input-output mechanisms. As technology becomes more integrated into daily life, there is a growing need for machines to understand and respond to human emotions in a meaningful way. This thesis focuses on the computational recognition of emotions through sentiment analysis—a subfield of Natural Language Processing (NLP)—and explores how it enhances the quality and effectiveness of human-computer interaction.

The thesis includes a preface on the impact of emotions on communication and judgment, demonstrating the importance of emotional intelligence in a wide range of systems such as e-learning platforms, virtual assistants, and customer care chatbots. It then covers the fundamental methods of sentiment analysis that would be necessary to classify user input as either good, negative, or neutral, while exploring rule-based methods, machine learning methods, and deep learning methods. One other purpose of these methods is to infer deeper emotional states like joy, anger, fear, or surprise. A large component of the thesis is assessing how well robots are able to identify the emotional tone of a user's text-based Realtime interactions, as well as how the incorporation of text, visual, and audio input-into multimodal sentiment analysis systems-can improve their ability to assess emotional tone. A significant portion of the thesis is devoted to examining current challenges in the field, including sarcasm detection, ambiguity in language, and cultural variances in emotional display. It also highlights directions for the future of research with a focus on context-aware systems that perform multimodal sentiment analysis, and posited potential solutions.

Overall, this study shows that sentiment analysis is essential to overcoming the emotional divide between people and computers and opening the door to more intelligent, sympathetic, and realistic human-computer interactions.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **SA** | Sentiment Analysis |
| **SER** | Speech Emotion Recognition |
| **NBC** | Naive Bayes Classifier |
| **KNN** | K- Nearest Neighbors |
| **CNN** | Convolutional Neural Network |
| **BERT** | Bidirectional Encoder Representations from Transfers |
| **LSTM** | Long short-term memory |
| **SVM** | Support Vector Machine |

# CHAPTER 1
# INTRODUCTION

Human attitudes, sentiments, judgments, inclinations, and other thoughts and emotions are fundamental to our decision-making process. Our actions are frequently impacted by the viewpoints and beliefs of others. This is due to the fact that humans are social creatures by nature, and as such, our capacity to trade and understand emotional cues influences how we interact and communicate. In the current digital era, this communication mostly takes place on websites like blogs, forums, YouTube, Facebook, and Twitter. These social media platforms have developed into effective means of sharing opinions, exchanging life stories, and influencing public opinion.

## 1.1 Problem statement

In today's digital world, understanding human emotions through both text and speech has become increasingly important, especially for building more intuitive and responsive systems. Social media platforms like Twitter serve as a rich source of public opinion, emotions, and sentiments. However, analysing tweets is challenging due to their short length, use of slang, abbreviations, and informal expressions. Natural Language Processing (NLP) offers powerful tools to process and interpret such data, enabling the extraction of meaningful emotional insights from tweets.

At the same time, speech emotion recognition is gaining prominence in areas like virtual assistants, mental health monitoring, and customer service. Unlike text, speech conveys emotion through pitch, tone, rhythm, and intensity, making it both valuable and complex to interpret. Deep learning models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Multi-Layer Perceptron's (MLP) can learn intricate patterns in audio signals and are widely used for emotion classification tasks.

Despite advancements, both tweet-based sentiment analysis and speech-based emotion recognition face issues like noisy data, class imbalance, and the subjective nature of emotions. This research aims to explore effective methodologies for analysing emotions from text using NLP techniques and from speech using hybrid deep learning models. By comparing and evaluating both approaches, the study seeks to contribute to the development of emotionally aware AI systems that can interact more naturally with humans.

## 1.2 Motivation

Sentiment analysis is also known as opinion mining, the (automated) identification, evaluation, and interpretation of emotional expressions. This field is an intersection of data mining, machine learning, and natural language processing (NLP). The key aim of sentiment analysis is to evaluate the emotional tone of a text and categorize it as neutral, negative, or positive. In more sophisticated methods, it can also identify particular emotions, such as happiness, anger, sadness, or sarcasm. With the increasing use of digital platforms, there is a huge amount of unstructured information coming from written comments, tweets, reviews, and other posts, each day. Such datapoints offer potential clues on sentiment, both individually and collectively. As a result, sentiment analysis is increasingly important in politics, economics, marketing, product development, and even healthcare. For example, political campaigns can use it to understand public reaction, and companies can use it to assess consumer satisfaction and brand perception. A variety of techniques may be used in the analysis, including rule-based techniques using existing sentiment lexicons, machine learning models using labelled datasets, and deep learning models which find patterns in large amounts of data.

Nonetheless, sentiment analysis has a number of difficulties. Accurate sentiment recognition is made challenging by the casual, grammatical, and frequently sarcastic language found on social media, which is also full of emoticons, slang, and acronyms. Furthermore, language and cultural disparities might result in a range of emotional responses, which makes interpretation much more challenging. Training datasets that are varied, well-structured, and well labelled are essential for obtaining significant results. Low-quality data might restrict the efficacy of sentiment analysis algorithms and result in inaccurate conclusions.

The ability to extract human emotion and opinion from digital text is made possible by sentiment analysis, to sum up. By bridging the gap between computer comprehension and human expression, it makes it possible for systems to react with more intelligence and empathy. Sentiment analysis will play an increasingly important role in market research, human-computer interaction, and decision-making processes as digital communication grows. Technology might become more socially and emotionally intelligent if this subject continues to grow.

## 1.3 Classification of Sentiment Analysis

Typically, SA involves a number of complex procedures, such as aspect or object-based extraction, sentiment categorization, subjectivity analysis, and opinion holder identification. When it comes to separating subjective judgments from objective statements in a document, subjectivity analysis is essential. Since they are not very useful for additional sentiment analysis, objective texts are usually removed at this point. The remaining information is then classified as either good, negative, or neutral based on an assessment of its sentiment polarity. Aspect or object-based extraction is particularly important among all activities since it is the foundation of SA. Identification of the opinion holder is also necessary in some situations, especially where knowing the view's creator offers substantial value.

### Subjectivity Classification

In sentiment analysis (SA), subjectivity categorization is an essential component. Differentiating between subjective and objective content in a text is its main objective . First, sentences or documents are classified into one of these two categories. For additional sentiment analysis, only subjective statements that convey feelings or views are taken into account. On the other hand, objective sentences are usually not included as they present factual data without adding any significant emotion. Take the statement, "I believe this Canon camera has a very large lens," for example. An opinion of a product feature is included in the statement, but it is not made apparent whether the attitude is favourable or unfavourable.

It might be taken to mean either annoyance (it doesn't fit in the bag) or advantage (higher photo quality), depending on the context. These ambiguous statements are frequently classified as neutral. The categorization of subjectivity as a stand-alone job has been extensively studied. Filtering away objective material and keeping just the subjective parts for sentiment analysis is the goal. To assist this, a variety of models have been created, including binary and Naïve Bayes classifiers, which are semi-supervised machine learning techniques. Unsupervised learning models have also been used more recently to improve the categorization procedure.

Subjective sentences are essential in SA because they capture human perspectives, evaluations, judgments, and personal beliefs. These are the expressions that sentiment analysis seeks to classify into positive, negative, or neutral categories based on the emotional content they convey.

## 1.4 Different Levels in Sentiment Analysis

Sentiment Analysis (SA) is commonly performed at three distinct levels: the document level, the sentence level, and the word or phrase level.

1. **Document Level**

At this level, the sentiment of an entire document is analysed as a whole. The goal is to determine the overall sentiment expressed in the text, assuming that it reflects the opinion of a single author or source. A common challenge here is **sentiment regression**, which aims to estimate the degree or intensity of sentiment—how strongly positive or negative the document is. Researchers have employed various approaches, including supervised learning models and linear combination methods that aggregate sentiment polarities throughout the text.
However, a major drawback of document-level analysis is that not every sentence in the text conveys an opinion. Many may be purely factual or objective. This can lead to less accurate sentiment classification. As a result, there has been a growing shift toward more fine-grained analysis at the sentence level, which allows for better filtering of objective content and more precise extraction of subjective, opinion-rich sentences.

**2. Sentence Level**

At the sentence level, sentiment analysis involves identifying whether each sentence in a text is subjective or objective. Once this distinction is made, only the subjective sentences are selected for further sentiment evaluation. Machine learning techniques are commonly employed to detect and classify these subjective sentences.

One method uses log probability scores and key root terms to assign sentiment values to each subjective sentence. Another method computes sentiment by averaging the emotional weight of individual words or terms in a sentence to determine the overall sentiment. Unique to these other methods, though, sentence-level sentiment analysis has some limitations. It can sometimes miss sentiment for sentences that are objective in their structure but still have sentiment that can be invisible. For instance: "This mug I just bought last week has cracks on the sides." While this statement provides factual information accurately, there is also quite a bit of subtle sentiment towards the durability of the product. In this example, because the sentence is objectively labelled, the sentiment may get missed. Text analysis at a smaller granularity—more specifically, at the word or phrase level—is necessary to capture hidden sentiment like the one in the example above. More information about this method is included in the next section.

### 3) Word/Phrase Level

The purpose of sentiment analysis at the word or phrase level is to analyze the most important elements of a text, which are single words and phrases. This method provides a detailed understanding of sentiment expressed in the language, so it is the most granular kind of sentiment analysis. Knowing the meaning and emotional valence of a word, sentiment analysis is able to consider each word separately, thereby allowing it to analyze more nuanced emotional expressions. The interest in the accuracy of this level of analysis has attracted a lot of attention from academic scholars and resulted in a sizable literature base. Early works explored the polarity of certain words and phrases, to assist with sentiment categorization on the level of sentences and documents. A prominent strategy is building sentiment lexicons - - a set of words associated with emotional values – either manually or by a computer. These identified words and phrases contained in lexicons often include adverbs (e.g., slowly, badly, elegantly), sentimental adjectives (e.g., beautiful, horrible, fantastic), and specific verbs (e.g., love, hate, like, detest). Sometimes, negative nouns, such as garbage or rubbish, also feature prominently.

## 1.5 Overview of Methods and Approaches

Among the various sentiment analysis (SA) methods there are two predominant methodologies in prevalent use. The first approach uses machine learning methods, and relies on identifying salient features that appropriately describe the polarity of sentiment in order to address SA problems. The main limitation of this method is having to rely on continuous observation due to requiring a manually annotated corpus for training. The second approach is a lexicon-based approach which focuses on linguistic features. As stated, this method begins with analyzing words or phrases that exhibit semantic polarity. There is a third approach that consists of components of both lexicon-based and machine learning-based approaches. This combined approach includes semi-supervised, or combined approach, and can contribute more value toward sentiment analysis.

### A. Machine Learning Approach

Two essential document collections are needed for machine learning (ML) text classification: the training collection and the test collection. While the test collection is used to evaluate the classifier's accuracy, the training collection teaches the classifier how to distinguish between distinct text characteristics.

There are many different ML techniques that can classify text as negative or positive. Some of the better strategies for classifying text are Support Vector Machines (SVM), Naive Bayes (NB) and Maximum Entropy (ME). Other approaches include ID3, Centroid Classifier, Winnow Classifier, K-Nearest Neighbor and Association Rules mining. The Naive Bayes (NB) classifier is typically used for text document classification. The Naive Bayes classifier is a probabilistic model that provides the probability of a particular category based on evaluating conditional probabilities of the terms and the categories in the document. As classifiers, Support Vector Machines (SVM) classifiers are mostly utilized with regards to solving problems where pattern recognition exists between two groups. The goal of SVM is to identify the hyperplane that optimally divides a data set into two groups. Initially, SVM was established for data that was linearly separable. SVM can also handle non-linear data if mapped in higher dimensions. SVM is popular among researchers because it is considered to be one of the best methods of classifying texts.

## B. DEEP LEARNING

Effective methods for identifying students who are at danger of failing a course include sentiment analysis and deep learning approaches. Convolutional neural networks (CNN) were employed in a study by Yu et al. (2018) to assess unstructured text comments from 181 undergraduate students as well as structured data, including grades and attendance. The Self-Assessment Manikin scale was used to manually classify the written response as either positive, negative, or neutral (Lang, 1980). The authors used this data to train CNN and support vector machine (SVM) models. In the fifth, seventh, and ninth weeks of the semester, the CNN model achieved F-measures of 0.78, 0.73, and 0.71, respectively,

Sutoyo et al. (2021) used a CNN model in a separate study to evaluate the effectiveness of lecturers based on student feedback from questionnaires. The results showed an accuracy of 87.95%, 87% precision, 78% recall, and an F1-score of 81%. It is possible that using attention in deep learning models could improve performance as they capture the sentiment of words with emotional variability. A multi-head attention fusion model was introduced by Sangeetha and Prabha (2020) to conduct sentiment analysis on student comments. Multi-head attention layers such as Glove and Cove operate using word and contextual embeddings to analyze a series of phrases simultaneously. The outputs of the multi-head attention layers are inputs to an LSTM model with dropout regularization to improve accuracy for the model. The study reported the performance of the fusion model increased accuracy in sentiment classification as compared with the multi-head attention model and LSTM model alone.

## C. Lexicon-Based Approach

The semantic orientation (SO) method is another approach used in sentiment analysis (SA). This technique operates with an unsupervised learning model, meaning that no initial training data is needed. It determines how closely a term aligns with either a positive or negative sentiment.

In the unsupervised learning approach, lexical rules are applied for sentiment classification tasks. A semi-supervised variant of this method uses WordNet as the key lexical resource [49], with a seed set derived from WordNet. The underlying principle is that words with similar orientations tend to share comparable glosses. To determine the semantic orientation of seed terms, a statistical approach using gloss classification is employed. A model using the k-means clustering technique is then applied, dividing documents into positive and negative categories. Afterward, a weighting method, such as term frequency-inverse document frequency (TF-IDF), is applied to the text, followed by a voting mechanism to improve the accuracy of results.

Some researchers have also developed sentiment lexicons, asserting that creating a manually crafted lexicon yields more precise results than an automatically generated one.

Moreover, previous research suggests that machine learning techniques generally perform better than sentiment orientation methods. This is because machine learning can handle multiple features to determine sentiment polarity. However, the sentiment orientation approach excels across different domains as it does not depend on domain-specific features. The challenge with sentiment orientation, however, lies in constructing a lexicon with accurate polarity, which is similar to the difficulty of selecting the right features in machine learning techniques.

### a. Corpus-Based Approach

The corpus-based approach focuses on identifying word co-occurrence patterns to assess sentiment. Different methods have been introduced for this task. For example, Peter Turney calculated the semantic orientation of a phrase by comparing its mutual information with the word "excellent" (representing positive sentiment) and subtracting its mutual information with the word "poor" (indicating negative sentiment). Similarly, Ellen Riloff and Janyce Wiebe used a bootstrapping technique to identify linguistically rich patterns of subjective expressions, which allowed them to differentiate between subjective and objective statements.

### b. Dictionary-Based Approach

The dictionary-based approach leverages synonyms, antonyms, and hierarchical structures found in resources like WordNet (or other lexicons containing sentiment data) to assess the sentiment of words. A more advanced version of WordNet, called SentiWordNet, serves as a specialized lexical resource for sentiment analysis, offering additional sentiment-related features. It assigns three sentiment scores—positive, negative, and objective—to each synset in WordNet. SentiWordNet has been widely used as the lexicon in recent sentiment classification research.

On the other hand, corpus-based techniques typically depend on a large corpus to collect the statistical data needed to determine the sentiment orientation of words or phrases. This can make them less efficient than dictionary-based methods. However, the effectiveness of dictionary-based approaches is heavily reliant on a comprehensive and accurate lexicon.

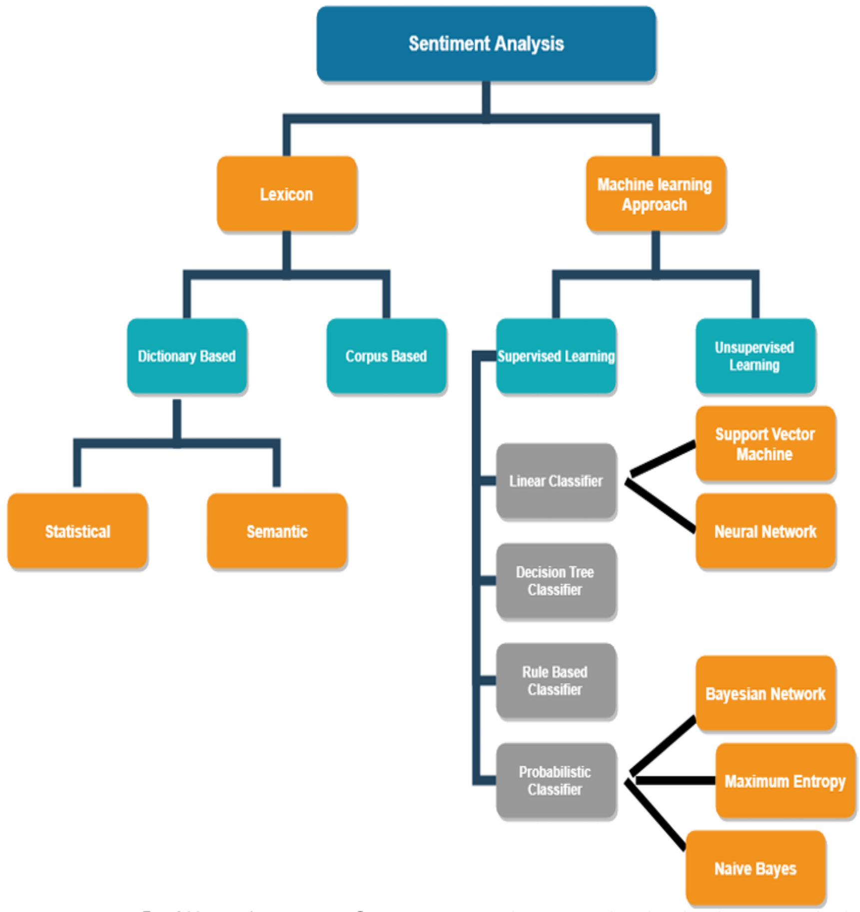


**Fig 1.1** Sentiment Analysis Techniques

## 1.6 NEED OF SENTIMENT ANALYSIS

### a. INDUSTRY EVOLUTION

In contrast to the entire collection of unstructured data, just a sufficient amount of data is needed in the industry. Nonetheless, the sentiment analysis that was conducted is helpful in identifying the crucial feature that can only be extracted from the data for industrial usage. Sentiment analysis will provide the industries a fantastic chance to add value and attract customers for themselves. This will help any industry that deals with consumers directly, including retail, dining establishments, entertainment, hospitality, and travel.

### b. RESEARCH

The growing need for research in assessment, evaluation, opinion, and categorization is a key factor fueling the expansion of Sentiment Analysis (SA). Technologies for opinion mining and sentiment analysis are advancing quickly, particularly in minimizing the human effort required for categorizing feedback. Furthermore, the field will leverage established computer science areas such as automated content analysis, voting advice applications, machine learning, text mining, and artificial intelligence.

### c. DECISION MAKING

In order to extract relevant information from blogs, diverse web applications, and social media platforms, a specific methodology is required for data analysis, yielding valuable insights. Conducting surveys regularly poses challenges for companies, thus emphasizing the importance of data analysis for identifying top-performing products based on user feedback, reviews, and recommendations. These reviews and opinions not only aid individuals in making informed decisions but also offer valuable insights for research and business endeavors.

### d. COMPREHENDING COMPLEXITY

As human language is getting very complex day by day so it has become difficult for the machine to be able to understand human language that can be expressed in the slangs, misspelling, nuances, and the cultural variation. Thus, there will be a need of system that will make better understanding between the human and the machine language.

### e. ONLINE MARKET

A major factor contributing to the increased demand for sentiment analysis is the rise of online marketing by businesses and organizations. These companies routinely track user opinions about their brand, product, or event through blogs and social media posts. As a result, sentiment analysis has also become a valuable tool for marketing.

## 1.7APPLICATION OF SA

Sentiment analysis, also known as opinion mining, finds applications across various domains due to its ability to extract and analyze subjective information from text data. Some of the key applications of sentiment analysis include:

- **Customer Feedback Analysis**: SA helps in analyzing surveys and getting reviews of customers which is later on used to improve customers need and to improve its experience.

- **Brand Monitoring:** Companies apply SA to monitor brand performance across various online platforms, helping them to understand public needs and their perception.

- **Market Research**: SA empowers brands to analyze market trends, consumer preferences, and competitor strategies by examining sentiments expressed in online discussions and reviews.

- **Financial Analysis**: SA is employed in finance to examine articles, posts and financial trends to understand people's inclination towards a brand or company.

- **Political Analysis:** SA It aids political analysts and campaigns in measuring public sentiment towards candidates, policies, and issues by analyzing media discussions and public speeches.

- **Customer Service Optimization:** SA is used to categorize and prioritize incoming requests by the customer to give them a overall good experience.

- **Healthcare**: Sentiment analysis is employed in healthcare to analyze patient feedback, reviews of healthcare providers, and social media discussions to assess patient satisfaction, identify areas for improvement, and enhance patient care.

- **Election Forecasting:** Sentiment analysis is used in political forecasting to analyze public sentiment towards political candidates and parties, helping to predict election outcomes more accurately.

## 1.8 Challenges in Text VS Speech Modalities

### Challenges in Text Modality

1. **Lack of Vocal and Visual Cues**
    - Emotions like sarcasm, irony, or excitement are harder to detect without tone, pitch, or facial expressions.
2. **Ambiguity and Context Dependence**
    - Words may have different meanings based on context (e.g., "sick" can mean ill or awesome).
3. **Use of Informal Language**
    - Emojis, slang, abbreviations, or misspellings (especially in social media) can hinder emotion detection accuracy.
4. **Limited Emotional Range**
    - Compared to speech, textual data often lacks subtle emotional variations unless enriched with strong sentiment markers.
5. **Sarcasm and Figurative Language**
    - Recognizing sarcasm or metaphor requires deep contextual and sometimes cultural understanding.

### Challenges in Speech Modality

1. **Noise and Audio Quality**
    - Background noise, poor microphone input, and channel distortions can degrade feature extraction.
2. **Speaker Variability**
    - Differences in age, gender, accent, and speaking style impact consistency and recognition performance.
3. **Real-time Processing Requirements**
    - Emotion recognition from live audio streams demands fast, efficient models with low latency.
4. **Emotion Overlap**
    - Some emotions have similar acoustic signatures (e.g., anger and excitement), making them difficult to differentiate.
5. **Multimodal Synchronization**
    - Combining speech with facial cues or body language for better accuracy is complex and resource-intensive.

# CHAPTER-2
# LITERATURE SURVEY

Sentiment analysis has long been a topic of interest in research. Contemporary studies concentrate on examining user-generated content from social media platforms such as Facebook, Twitter, and Amazon. These investigations predominantly use machine learning methods to identify the sentiment polarity of texts and to assess whether a given expression supports or opposes a particular viewpoint. This chapter presents several research findings that further enhance our knowledge of sentiment analysis

**P. Pang, L. Lee, S. Vaithyanathan et al**

They were the first to investigate sentiment analysis by aiming to classify text based on its sentiment rather than its subject—for example, identifying whether a movie review expresses a positive or negative opinion. By applying machine learning algorithms to a collection of movie reviews, they discovered that these methods outperformed those based on manually crafted rules. The techniques used included Naïve Bayes, Maximum Entropy, and Support Vector Machines. Their work also concluded that sentiment classification is inherently complex and that supervised machine learning algorithms are vital for effective analysis

**P. Pang, L. Lee et al**

Rapid progress in data mining and sentiment extraction focuses on applying computational power to overcome challenges in opinion mining and detecting subjectivity in texts. This has led to the development of new systems built with diverse languages and commands that prioritize opinion mining, emphasizing real-time processing and ongoing research. Their survey reviews the methodologies and approaches employed in direct opinion mining, highlighting the most effective ones. They concentrate on features that address emerging challenges in sentiment analysis, comparing these advanced techniques with traditional fact-based methods.

**E. Loper, S. Bird et al**

The Natural Language Toolkit (NLTK) library includes a large number of software modules, a large number of structured files, tutorials, problem sets, statistical functions, machine learning classifiers, and courseware for computational linguistics. Its main objective is to support natural language processing by analysing human language data. With the corpora that NLTK offers for classifier training, developers may modify existing components and construct new ones, as well as design more organized systems that generate more sophisticated outputs from the dataset.

**H. Wang, D. Can, F. Bar, S. Narayana et a**

By establishing a connection between the opinions shared on Twitter and the larger election events, they looked into how sentiment analysis may affect how these public events are understood. They also showed how much faster this real-time sentiment analysis is than traditional content analysis, which usually takes days or weeks to finish. With real-time findings, the system they created continually analysed Twitter data about the election, candidates, and promotions. Researchers, legislators, and the media may measure public opinion more quickly and effectively with this strategy.

**M. Yu, M. Zhou, X. Liu, T. Zhao et al**

State-of-the-art methods only consider tweets when they are utilized for categorization. Since these methods categorize primarily on the most recent tweet, they exclude related tweets. However, analysing a current tweet alone for sentiment analysis is insufficient due to the fact that tweets are typically brief and vague. They suggest a method to enhance target-dependent sentiment categorization on Twitter by:

1) Including target-dependent characteristics, and

2) Considering trending tweets.

These recent developments significantly enhance the effectiveness and performance of target-dependent sentiment categorization, based on their experimental findings.

## 2.1 OVERVIEW OF EXISTING SYSTEM

Considerable research has been conducted on sentiment analysis from transcribed speech. However, there has been less emphasis on analysing speech sentiment based purely on acoustic characteristics. It has been observed that the acoustic qualities of speech change depending on the emotional state of the speaker, which suggests that a speaker might express a positive sentiment verbally while feeling differently emotionally.

Researchers have investigated a number of classification techniques, including Support Vector Machines (SVM), K-Nearest Neighbours (KNN), Multiple Likelihood Bayes Classifier (MLC), Hidden Markov Model (HMM), Gaussian Mixture Model (GMM), Neural Networks (NN), and Kernel Regression. An algorithm that analysed vocal characteristics such pitch, timbre, sound pressure level (SPL), and speech intervals was created in order to differentiate between emotions such as normal, furious, and terrified. It was made simpler to distinguish between different emotional states by deriving numerical values for these voice characteristics using programs like Wavepad and MATLAB.

pyAudioAnalysi is one of the more well-known Python packages created especially for processing audio streams. This package makes feature extraction, segmentation, regression, classification (including SVM and KNN), and visualization possible. Particularly noteworthy are pyAudioAnalysis's thorough feature extraction connected to machine learning elements and baseline methods designed for audio analysis jobs.

Sentiment analysis from speech presents several challenges, including background noise, foreign accents, real-time speech generation, and the wide variety of speech topics. Two main approaches to speech analysis are commonly employed: one involves analysing the transcribed text while ignoring the acoustic features to identify sentiment, and the other preserves the acoustic characteristics, focusing on the underlying sentiment within the sound. This study utilizes three distinct datasets, while most research typically relies on a single data source.

## 2.2 RELATED WORK

### 2.2.1 SA of Twitter Tweets

**A. Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques**

This study presented a machine learning (ML) approach to the analysis of sentiment tweets in the Twitter dataset. The TF-IDF methodology was utilized for feature extraction in order to pick effective features for the two main components of the proposed method, feature extraction and classification. As classifiers, a variety of classification techniques have been successfully tried. To assess the approach, the US airline sentiment Twitter dataset was employed. In conclusion, SVM produces the best results, with an improved F1-score of 87.85% and 83.74% accuracy and 84% precision. Future research will examine more feature extraction methods.

**B. Sentiment Analysis Framework of Twitter Data Using Classification**

Given that the machine learning algorithms were run on an untested dataset, language hurdles have also been shown to affect the programs' current viability and future scope. Random Forest, Support Vector Machine, and Naïve Bayes classification techniques were used in the investigation. Python was utilized to carry out this experiment, and Excel was utilized to further assess and plot some of the findings. In order to avoid mistakes while assessing the models' accuracy and precision, a test set has been manually generated because it is impossible to determine the sentiment of the tweets.

**C. Sentiment Analysis with NLP on Twitter Data**

They have created a sentiment analysis model in this work that enables real-time processing of Twitter API streaming feeds and the classification of their polarity, offering useful information about users and the industry. NLTK may make use of our constructed classifier as a data analysis tool. Because our suggested method outperforms all current models in terms of accuracy, we can generally use it to analyse sentiment for any gadget, public person, or sports team.

**D. Sentiments Analysis of Twitter Data Using Data Mining**

The experimental findings show that data mining classifiers are a suitable option for sentiment prediction utilizing tweeter data. In a test, knearest neighbor (IBK) performs better than the three classifiers—Random Forest, BaysNet, and Naive Baysein. Additionally, Random Forest has high forecast accuracy. For sentiment predictions of tweets, an ensemble of classifiers is not necessary because a single classifier (i.e., knearest neighbor) provides a higher accuracy than any combination of ensemble classifiers.

**E. Twitter Sentiment Analysis Using Lexical or Rule Based Approach: A Case Study**

It can be concluded that lexical or rule-based approaches (unsupervised learning) often lack accuracy due to the frequent use of slang in social media texts. Therefore, after categorizing raw data into positive, negative, or neutral using manual methods or built-in libraries, it is necessary to apply supervised machine learning algorithms for more precise results. The structured data is then split into training and testing sets before being used to train the classification model. Supervised methods are considered more specialized than unsupervised techniques relying on built-in libraries. To achieve greater accuracy, further development of approaches that integrate both supervised and unsupervised methods is required.

**F. Unveiling Twitter Sentiments: Analyzing Emotions and Opinions through Sentiment Analysis of Twitter Dataset**

Finally, the study investigated how well machine learning-based sentiment analysis models performed on data from Twitter. Vader and Roberta, two well-known sentiment analysis algorithms, were used to compare the dataset's outcomes. Although there were some variations in their trial findings, Vader and Roberta both did a good job of correctly determining the sentiment of the tweets in our sample. More specifically, compared to Vader, Roberta tended to assign a lower negative score and a higher positive score. Roberta is a more sophisticated deep learning model, so it can better grasp sentiment and linguistic subtleties. Additionally, they merged and contrasted the outcomes of the Roberta and Vader models.

### 2.2.2 Speech Emotion Recognition System

In this section, we will examine prior research and studies that are relevant to the scope of our project. A significant amount of work has been conducted in the field of speech emotion detection, mainly focusing on acoustic features of sound. During the project preparation, we thoroughly reviewed various research materials and models, gaining valuable insights and inspiration. Below, we will discuss these influential studies and findings.

The task involves developing a system designed to analyze and classify speech signals within audio samples, aiming to detect and identify emotions expressed within these signals. The emotions will be classified in the classes such as male happy, male sad, female happy, female sad, male angry, female angry, etc. This system has diverse potential applications, including usage in interactive voice-based assistants and the analysis of conversations between callers and agents. Speech Data Preprocessing In speech, three primary classes of features contribute to its characterization: lexical features (pertaining to the vocabulary employed), visual features (indicative of the speaker's expressions), and acoustic features (encompassing sound properties such as pitch, tone, and jitter, among others). Analyzing one or more of these features can effectively address the challenge of speech emotion recognition. In this project, the focus lies on the extraction of Mel Frequency Cepstral Coefficients (MFCC) from the audio data. These coefficients serve as representations of the audio's pitch, perceived frequency, and tonal qualities. The extraction process was implemented in Python using the Librosa package. Librosa: It is a Python package dedicated to audio analysis, Librosa encompasses various submodules.

These include functionalities for beat estimation, core audio operations (e.g., loading audio, computing spectrograms), visualization using matplotlib, and extraction of diverse audio features like chromagrams, Mel spectrograms, and MFCCs. MFCC (Mel Frequency Cepstral Coefficients) : In classical time signal analysis, recurring elements such as echoes manifest as clear peaks within the frequency spectrum. This spectrum is obtained by applying the Fourier transform to the temporal signal. Cepstral features, such as Mel-Frequency Cepstral Coefficients (MFCCs), are obtained by performing a Fourier Transform on a spectrogram. MFCCs possess a unique characteristic as they are computed on a Mel scale, which adjusts frequencies to align more closely 9 with human auditory perception. This scaling ensures a better correspondence to the frequencies perceivable by the human ear. Importantly, MFCCs effectively capture the temporal power spectrum of speech signals, portraying the attributes of the vocal tract. Models Involved This phase encompasses several steps: shuffling the data, dividing it into training and testing sets, and constructing various models for training. The intention is to assess and select the most suitable model based on its performance with the available dataset. The models under consideration involve a MLP model, LSTM model, and CNN models. Outcome and Evaluation The trained model will undergo evaluation using the test dataset, and its performance will be assessed through a confusion matrix. The model that exhibits the most favorable performance metrics will be selected as the final model for emotion detection in human speech. The model's prediction capability involves taking any recorded input audio and determining the embedded emotion from the following predefined classes:

"0 - female_angry 1 - female_calm 2 - female_fearful 3 - female_happy 4 - female_sad 5 - male_angry 6 - male_calm 7 - male_fearful 8 - male_happy 9 - male_sad”

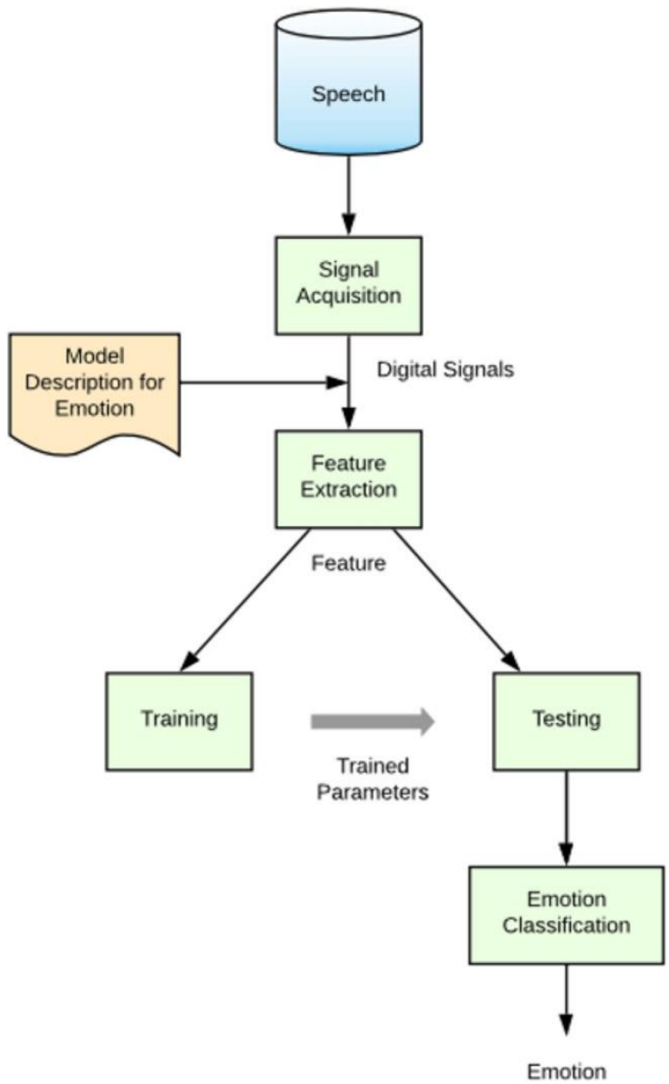


**Fig 2.1** Block diagram of SER

# CHAPTER-3
# RESEARCH METHODOLOGY

Contrary to appearances, gathering data is not an easy undertaking. For the purpose of gathering data, many choices must be taken. For this project, I have collected datasets for sentiment analysis on Twitter, testing, and training. The collection, storing, processing, and classification of data will all be covered in this chapter. Let's first talk about our suggested architecture before moving on to these procedures and other datasets.

## 3.1 DATASET PROFILE

### CASE-I Text Based (SENTIMENT ANALYSIS OF TWITTER TWEETS)

**Twitter's API** enables users to gather tweets. Twitter offers two different types of APIs: streaming and REST. These differ from one another in the following: In contrast to streaming APIs, which offer real-time tweets and connections for extended periods of time, REST APIs only allow connections for brief periods of time and can only gather a limited amount of data at once. We do our analysis using the Streaming API. We require a long-lasting connection and an unlimited data rate in order to gather a big number of tweets.

We may get data from our Twitter accounts and utilize it for our own purposes thanks to the platform that Twitter provides. We must use our Twitter login information to access the dev.twitter.com website in order to do this. Using the information provided, we first develop an application on this website that will be used to stream tweets. We are able to obtain the customer key, customer secret key, access token key, and access secret key after our API is built. Users that wish to access Twitter data must authenticate using these keys.

### CASE-II Speech Based (SPEECH EMOTION RECOGNITION SYSTEM)

The dataset used for the project is divided into two groups:

1. "**The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) data set**".
2. "**Surrey Audio-Visual Expressed Emotion (SAVEE) dataset**"

The Ryerson Audio-Visual Database of Emotional Speech and Song - RAVDESS data set: The roughly 1500 audio files in this dataset were provided by 24 different actors, 12 of whom are male speakers and the remaining 12 are female. The eight distinct emotions that these actors captured in short audio clips are labeled numerically as follows: 1 denotes neutrality, 2 calmness, 3 happiness, 4 sadness, 5 anger, 6 fear, 7 disgust, and 8 astonishment. Every audio file has a unique name, and the seventh character always denotes the particular emotion being conveyed.

Visualizing the Dataset: We conducted an analysis on an audio file to extract its features, visualizing both its waveform and spectrogram.

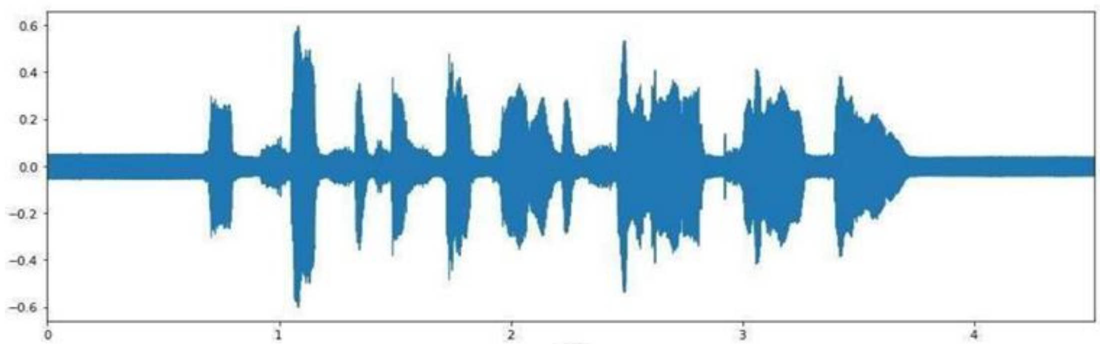


**Fig 3.1** Sample Audio Waveform

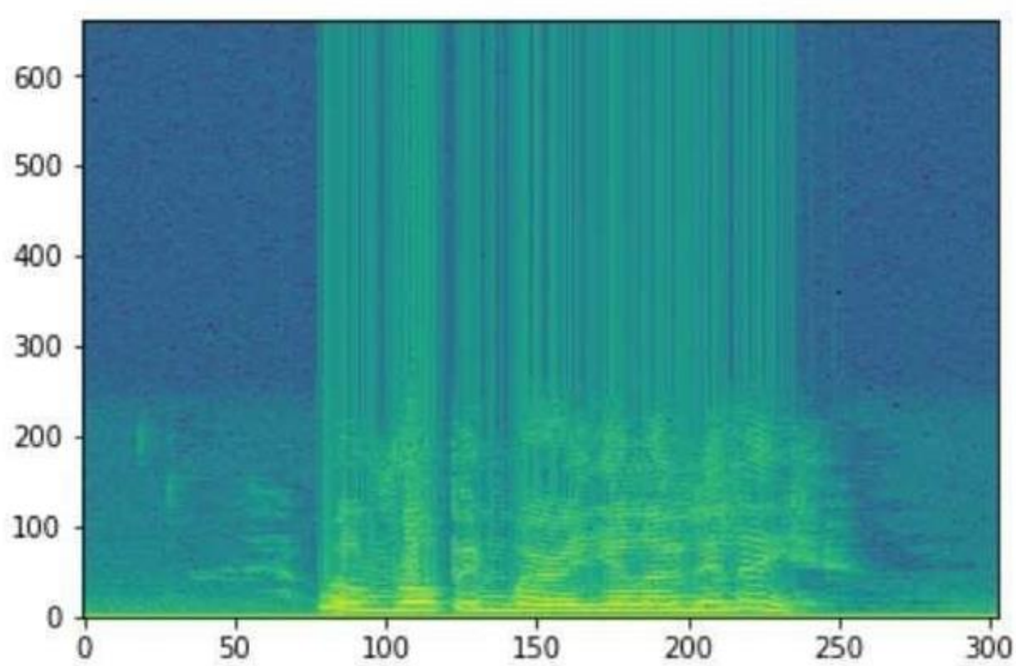


**Fig 3.2** Sample Audio Spectrogram

Surrey Audio-Visual Expressed Emotion - SAVEE dataset: This dataset comprises roughly 500 audio files recorded by four distinct male actors. Each file name begins with two characters that denote the specific emotions conveyed within.

## 3.2 FRAMEWORK OVERVIEW

### 3.2.1 Convolutional Neural Network (CNN):

Convolutional Neural Networks (CNNs) are a subset of deep neural networks that are particularly useful for the processing of visual images. To distinguish CNNs from other networks, CNNs are special types of shift-invariant networks due to the presence of convolutional kernels or filters in their architecture. In addition, through collective-weight architectures CNNs can successfully identify visual input. When applied to input features, these filters the filters slide across completely the input producing feature maps, which vary exceedingly where the output of our feature map is determined as the translation-equivariant.

CNNs are widely used in applications such as natural language processing, image/video recognition, image classification, medical image analysis, and others. MLPs are thought of as a slightly more complex architecture derived from CNNs. Overfitting will be arise from classical networks (e.g., MLPs) as they have complete connections. Though CNNs contain these architectures in favor of larger constructs or patterns through hierarchical processing of data. In short, CNNs will arrange smaller and simpler patterns in input data with subsequent use of filters to produce larger and more complex patterns in the model outputs. This hierarchical approach provides an element of regularization, while extracting meaningful features from the input, so that CNNs provide practical features extraction and pattern recognition at lower levels of connectivity and complexities.

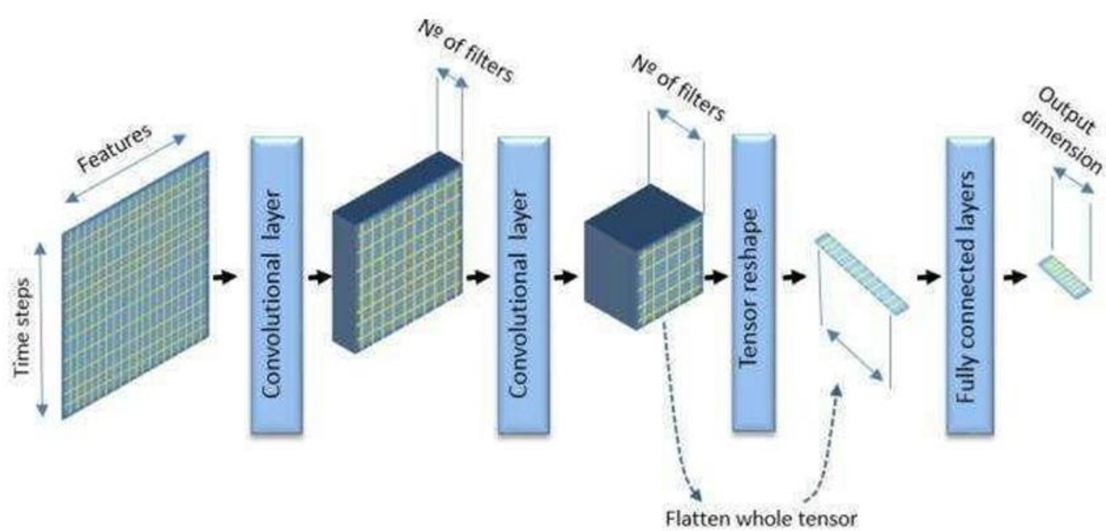


**Fig 3.3** CNN model representation

**3.2.2 Naive Bayes Classifier**:

The **Naive Bayes classifier** is based on **Bayes' Theorem**, a mathematical formula used to determine the probability of a class given some observed data. It's called "naive" because it assumes that all features (words in this case) are **independent** of each other — which is rarely true in real text but works well in practice.

We compare:

- P (Positive Tweet) P (\text {Positive} | \text {Tweet}) P (Positive Tweet)
- P (Negative Tweet) P (\text {Negative} | \text {Tweet}) P (Negative Tweet)

The tweet is classified into the class with the higher probability.

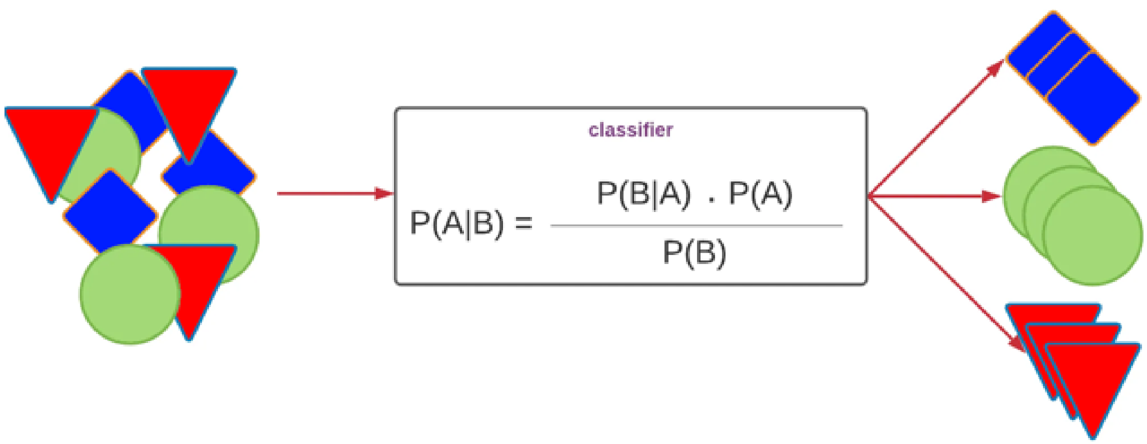


**Fig 3.4** Naïve Baye’s model representation

## 3.3 IMPLEMENTATION FLOW

### 3.3.1 Text Based

We want to accomplish sentiment analysis using data from Twitter. We intend to construct a classifier that incorporates various machine learning classifiers. The procedures depicted in Figure 3.1 will be followed once our classifier is prepared and trained.

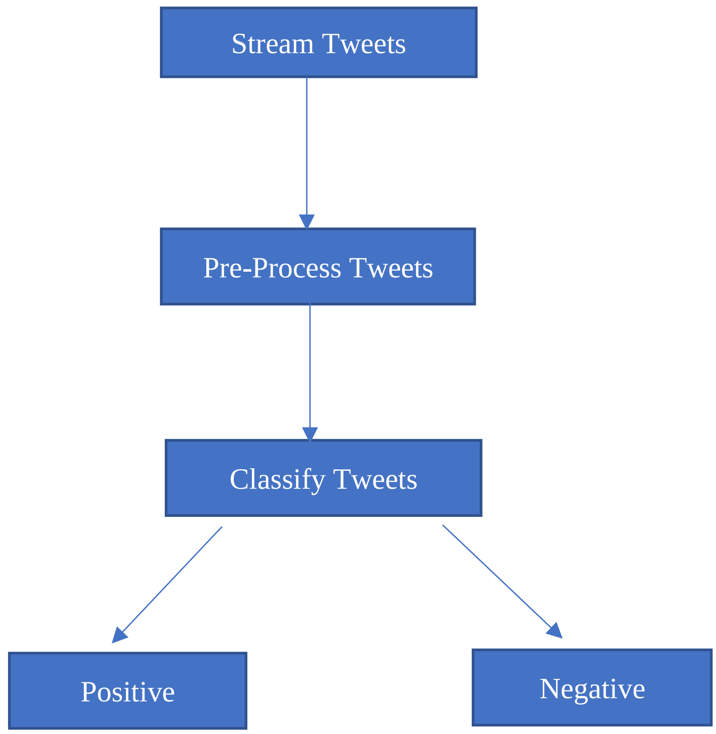


**Fig 3.5** Using a built classifier to classify tweets

#### 3.3.1.1 TOKENIZATION:

Regarding Natural Language Processing (NLP) and machine learning, tokenization is the act of breaking up a text sequence into smaller units called tokens. The length of these tokens might range from words to characters.

A textual sequence of letters that functions as a unit is called a token. The tokens can be text, emojis, hashtags, URLs, or even single letters, depending on how you make them.

Splitting the text according to punctuation and whitespace is a simple method of dividing language into tokens.

### 3.3.1.2 NORMILIZATION:

One of the most important stages in natural language processing (NLP) is text normalization. In order to make text data consistent and useful for various NLP tasks, it must be cleaned and pre-processed. Numerous methods are used in the process, including lemmatization, stemming, stop word removal, punctuation removal, and case normalization.

### 3.3.1.3 DATA-PREPROCESSING:

Data from Twitter is unsuitable for feature extraction. The majority of tweets include a message, username, time stamps, URLs, special characters, stop words, emojis, abbreviations, hash tags, and other information. For this reason, we pre-process this data using several NLTK functions to make it suitable for mining. Following the extraction of the tweet's primary message, pre-processing involves eliminating any blank spaces, stop words (such as is, a, the, he, they, etc.), hash tags, repetitive words, URLs, and other unnecessary information. Next, we add the appropriate meanings to all emoticons and abbreviations, such as :-), =D, =), LOL, Rolf, etc., and substitute them with happy or laugh. After we finish processing it, we have a finished tweet that can be sent to the classifier to produce the desired outcomes. In Table 3.2, a processed tweet example is displayed.

**Table 3.1** sampled tweet vs processed tweet

| TWEET TYPE | RESULT |
|---|---|
| Original Tweet | @XYZ I think Kejriwal is a habitual liar, even where he doesn’t need to lie, he tells a lie >: p#AAP |
| Processed Tweet | Think, habit, lie, even, doesn’t, lie, angry |

Given that tweets contain a number of grammatical elements that could not be helpful for analysis, cleaning Twitter data is essential. Pre-processing is carried out such that the data is only expressed in terms of phrases that make classification simple.

We write a Python code that defines a function that will be used to get processed tweets. The following functions are accomplished with this code:

- ∑ remove quotes
- ∑ remove @
- ∑ remove URL (Uniform resource locator)
- ∑ remove RT (Re-Tweet)
- ∑ remove Emoticons
- ∑ remove duplicates
- ∑ remove #
- ∑ remove stop words - Eliminate all stop words that don't add sense to the categorization, such as a, he, the, and so forth.

Both the categories of material included in tweets and the actions taken on them are displayed in Table 3.3. Here are some examples of tweets that are clean.

**Table 3.2** Removed and modified content

| CONTENT | ACTION |
|---|---|
| Punctuation (! ? , . " : ; ) | Removed |
| #word | Removed #word |
| @any_user | Remove @any_user or replaced with "AT_USER" and then added in stop words. |
| Uppercase characters | Lowercase all content |
| URLs and web links | Remove URLs or replaced with "URL" and then added in stop words |
| Number | Removed |
| Word not starting with alphabets | Removed |
| All Word | Stemmed all word (Converted into simple form) |
| Stop words | Removed |
| Emoticons | Replaced with respective meaning |
| White spaces | Removed |

**Table 3.3** Sample cleaned data

| Raw data | Clean data |
|---|---|
| @jackstenhouse69 I really liked it, in my opinion it def is :) | Really, liked, opinion, def |
| :( \u201c@EW: How awful. Police: Driver kills 2, injures 23 at #SXSW http://t.co/8GmFiOuZbS\u201d | Sad, awful, police, driver, kills, injures |

Determining word density in Natural Language Processing (NLP) involves calculating the frequency or occurrence of words within a given text or corpus.

Word density can provide insights into the importance, relevance, or patterns of specific words in the text.

### 3.3.1.4 Preparation of Data Model

We want to see positive and negative sentiments therefore we will convert the tokens into dictionary. Splitting the dataset for training and testing the model. The next thing we need to do is to train the data for the Naïve Bayes' classifier class.

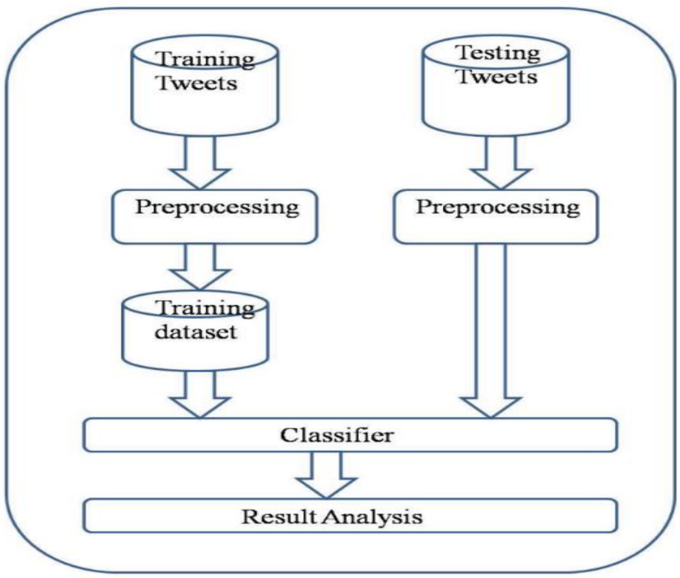


**Fig 3.6** Methodology

Evaluation metrics:
To assess how well your classifier is performing, use measures like accuracy, precision, recall, F1-score, and confusion matrix.

Cross-validation:
Perform k-fold cross-validation to get a better estimate of the model's performance and to ensure its generalization to unseen data.

Overfitting/Underfitting:
Check for signs of overfitting or underfitting by comparing the performance on the training set and the test set.

### 3.3.2 Speech Based

#### 3.3.2.1 Setting the labels for dataset:

The audio file within dataset is tagged with label denoting the embedded emotion. In the RAVDESS dataset, the 7th character in the file name consistently represents the specific emotion, while in the SAVEE dataset, the initial two characters indicate the portrayed emotions.

**Table 3.4** Labelling dataset with their embedded emotions

| | 0 |
|---|---|
| 0 | Male_calm |
| 1 | Female_calm |
| 2 | Male_calm |
| 3 | Female_calm |
| 4 | Male_calm |
| 5 | Female_calm |
| 6 | Male_calm |
| 7 | Female_calm |
| 8 | Male_calm |
| 9 | Female_calm |

#### 3.3.2.2 Feature Extraction:

In order to help our model, recognize differences between the audio files, the next step concentrates on feature extraction from the files. We use the Librosa package, which is based on Python and is known for its efficiency in audio analysis, for this feature extraction operation. There are certain factors to take into account in this process: All audio files are accurately timed for a period of three seconds during the feature extraction process. By guaranteeing a same quantity of features throughout the dataset, this standardized length encourages uniformity in model training. Higher Sampling Rate: The sampling rate for each file has been doubled while the frequency of sampling has remained constant. Enhancing the dataset's feature variety is the goal of this modification, which is especially helpful when working with smaller datasets. The increased feature count facilitates improved audio categorization.

**MFCC (Mel Frequency Cepstral Coefficients)**: -

In classical time signal analysis, echoes are simply recognized as attributed peaks in the frequency spectrum. This spectrum is generated by using a Fourier transform on the temporal signal. By performing a Fourier Transform on a spectrogram, you can provide cepstral features like the Mel-Frequency Cepstral Coefficients (MFCCs). MFCCs are unique in that they are estimated on a Mel scale, which adjusts frequency to more closely match the perceptual domain for human hearing. The Mel scale allows the frequencies to match more precisely to what the human ear can hear. It's important to note that MFCCs track the envelope of the temporal power spectrum in speech signals, characterizing properties of the vocal tract.

```
In [60]: train[255:265]
Out[60]:
```

| 4 | 5 | 6 | 7 | 8 | 9 | ... | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 0 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 582 | 0.243815 | 0.234133 | 0.220812 | 0.222221 | 0.232087 | ... | 0.248799 | 0.253912 | 0.260256 | 0.257698 | 0.258209 | 0.256242 | 0.255648 | 0.255648 | 0.255701 | angry |
| 521 | 0.285065 | 0.291352 | 0.303514 | 0.308232 | 0.328804 | ... | 0.234485 | 0.228035 | 0.216631 | 0.214859 | 0.212437 | 0.213037 | 0.218348 | 0.223208 | 0.224450 | fearful |
| 765 | 0.108862 | 0.103840 | 0.101478 | 0.107730 | 0.103912 | ... | 0.066940 | 0.036635 | 0.027208 | 0.036532 | 0.053178 | 0.065569 | 0.057186 | 0.039764 | 0.021314 | angry |
| 141 | 0.074467 | 0.089486 | 0.088280 | 0.092139 | 0.093846 | ... | 0.054423 | 0.053604 | 0.055540 | 0.058426 | 0.060729 | 0.068808 | 0.088886 | 0.098216 | 0.090357 | sad |
| 724 | 0.281591 | 0.296421 | 0.285957 | 0.260214 | 0.257237 | ... | 0.299710 | 0.291853 | 0.291916 | 0.299710 | 0.299710 | 0.299710 | 0.287766 | 0.252755 | 0.243608 | happy |
| 779 | 0.330779 | 0.330779 | 0.330779 | 0.330779 | 0.330779 | ... | 0.288739 | 0.287423 | 0.283312 | 0.291878 | 0.305482 | 0.321055 | 0.327999 | 0.301280 | 0.300456 | calm |
| 433 | 0.169379 | 0.171645 | 0.179289 | 0.190308 | 0.182795 | ... | 0.149075 | 0.147707 | 0.159900 | 0.184663 | 0.187635 | 0.168762 | 0.149145 | 0.130382 | 0.120786 | neutral |
| 036 | 0.238554 | 0.242728 | 0.229463 | 0.228398 | 0.243454 | ... | 0.223064 | 0.207814 | 0.210600 | 0.210909 | 0.202713 | 0.192792 | 0.192630 | 0.195298 | 0.187149 | happy |
| 079 | 0.326079 | 0.305091 | 0.284397 | 0.274060 | 0.266039 | ... | 0.156601 | 0.185422 | 0.202734 | 0.204833 | 0.213753 | 0.221158 | 0.222267 | 0.185138 | 0.151496 | sad |
| 975 | 0.172604 | 0.173216 | 0.167372 | 0.168891 | 0.178888 | ... | 0.205757 | 0.200951 | 0.197044 | 0.193599 | 0.208915 | 0.228052 | 0.219472 | 0.205900 | 0.201549 | surprised |

**Fig 3.7** MFCC features dataframe

**Model Training and Prediction**

A convolutional neural network (CNN), a long short-term memory (LSTM) network, and a multi-layer perceptron (MLP) are among the models used.

We next used our test dataset to predict emotions after training our model. We compared the expected and actual results using accuracy measures. During this process, a confusion matrix comprising metrics such as false positive (FP), false negative (FN), true positive (TP), and true negative (TN) was created.

# CHAPTER 4
# RESULT AND DISCUSSION

We will demonstrate a number of outcomes from our implementation in this chapter.

## 4.1. Analysis of Text Based Data

### CLASSIFIER ACCURACY FOR THIS TRAINING DATA

```
from nltk import classify
from nltk import NaiveBayesClassifier
classifier = NaiveBayesClassifier.train(train_data)

print("Accuracy is:", classify.accuracy(classifier, test_data))

print(classifier.show_most_informative_features(10))
```
[18] Python

```
Accuracy is: 0.998
Most Informative Features
                      :( = True           Negati : Positi =   2058.3 : 1.0
                      :) = True           Positi : Negati =   1643.0 : 1.0
                follower = True           Positi : Negati =     23.4 : 1.0
                     sad = True           Negati : Positi =     23.0 : 1.0
                     bam = True           Positi : Negati =     20.3 : 1.0
                    glad = True           Positi : Negati =     20.3 : 1.0
                     x15 = True           Negati : Positi =     18.3 : 1.0
                  arrive = True           Positi : Negati =     14.4 : 1.0
                     via = True           Positi : Negati =     13.4 : 1.0
```

**Figure 4.1** classifier accuracy for training data

Fig 4.1 demonstrates that our classifier is providing an average **accuracy of 99.8%.** As a result, our constructed classifier is completely trained and prepared for sentiment analysis of Twitter data.

### Twitter Analysis for Selected Phrases

1. **"I LOVE MY MOTHER"**

   **O/P: POSITIVE**

```
from nltk.tokenize import word_tokenize

custom_tweet = "I love my mother"

custom_tokens = remove_noise(word_tokenize(custom_tweet))

print(classifier.classify(dict([token, True] for token in custom_tokens)))
```

```
Positive
```

**Fig 4.2** Text Output1

2. **"CONGRATS SPORTSTAR ON YOUR 7$^{TH}$ GOALFROM LAST SEASON WINNING GIAL OF THE YEAR"**

   **O/P: POSITIVE**

```
custom_tweet = 'Congrats #SportStar on your 7th best goal from last season winning goal of the year :) #Baller #Topbin #
custom_tokens = remove_noise(word_tokenize(custom_tweet))
print(classifier.classify(dict([token, True] for token in custom_tokens)))
```

```
Positive
```

**Fig 4.3** Text Output2

3. **"SORRY, YOU ARE NOT ELIGIBLE"**

   **O/P: NEGATIVE**

```
custom_tweet = 'sorry,you are not eligible'
custom_tokens = remove_noise(word_tokenize(custom_tweet))
print(classifier.classify(dict([token, True] for token in custom_tokens)))
```

```
Negative
```

**Fig 4.4** Text Output3

## 4.2 Analysis of Speech Based Data

### 4.2.1 Result Analysis for SAVEE Dataset

Test Accuracy: 97.92%

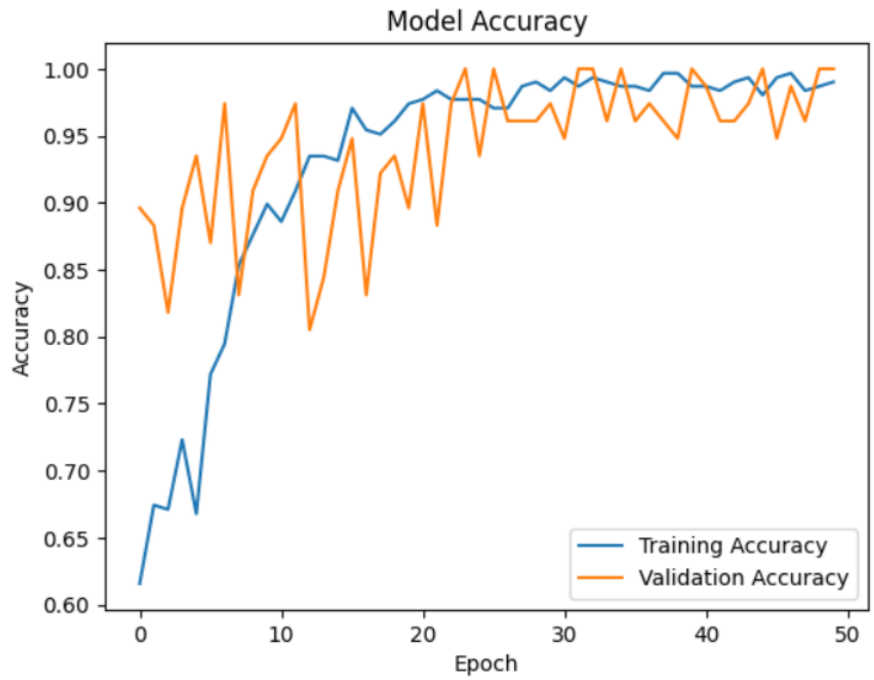


**Fig 4.5** CNN accuracy graph for SAVEE

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| disgust | 0.92 | 1.00 | 0.96 | 24 |
| neutral | 1.00 | 0.97 | 0.99 | 72 |
| accuracy | | | 0.98 | 96 |
| macro avg | 0.96 | 0.99 | 0.97 | 96 |
| weighted avg | 0.98 | 0.98 | 0.98 | 96 |

**Fig 4.6** SAVEE Classification Report

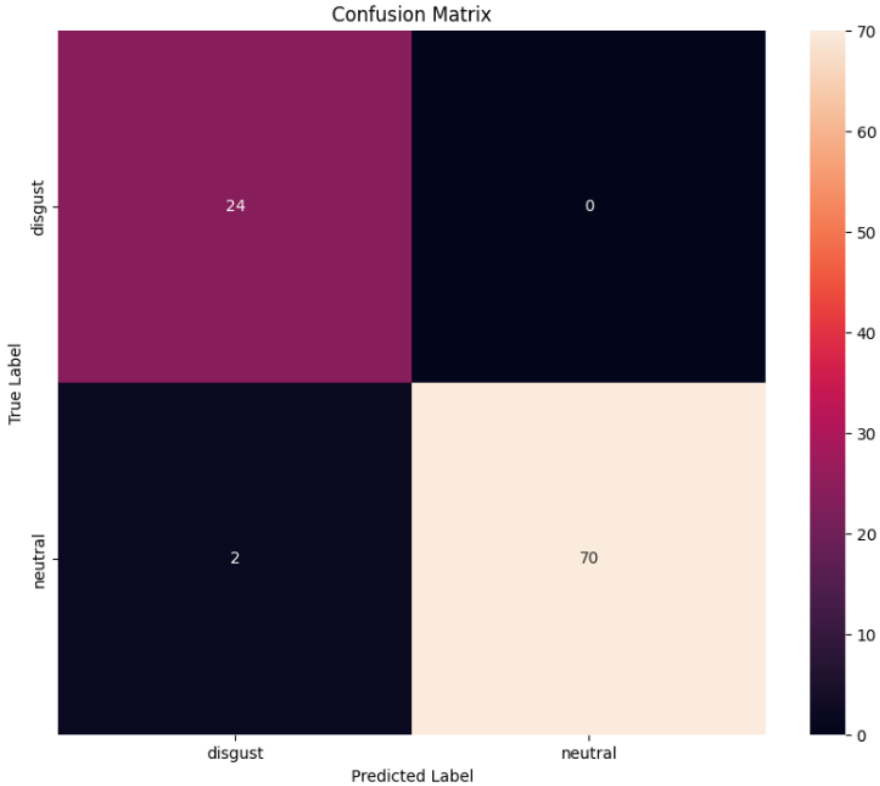


**Fig 4.7** SAVEE Confusion Matrix

### 4.2.2 Result Analysis for RAVDESS Dataset

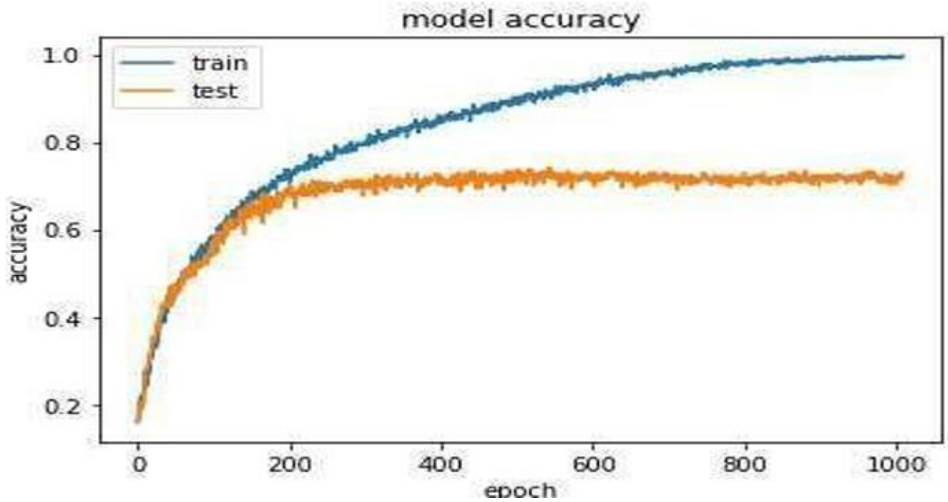


**Fig 4.8** CNN accuracy graph for RAVDESS

Confusion Matrix Values:

Predicted →

| Actual ↓ | Neu | Calm | Hap | Sad | Ang | Fear | Dis | Sur |
|---|---|---|---|---|---|---|---|---|
| Neutral | 112 | 15 | 9 | 12 | 7 | 10 | 8 | 7 |
| Calm | 6 | 153 | 4 | 5 | 3 | 4 | 3 | 2 |
| Happy | 11 | 7 | 104 | 18 | 12 | 15 | 9 | 4 |
| Sad | 14 | 9 | 22 | 95 | 13 | 16 | 7 | 4 |
| Angry | 5 | 4 | 8 | 9 | 133 | 12 | 5 | 4 |
| Fearful | 9 | 6 | 14 | 17 | 11 | 113 | 6 | 4 |
| Disgust | 8 | 5 | 7 | 10 | 9 | 8 | 106 | 7 |
| Surprised | 6 | 3 | 5 | 6 | 5 | 5 | 4 | 122 |

**Fig 4.9** RAVDESS Classification Report

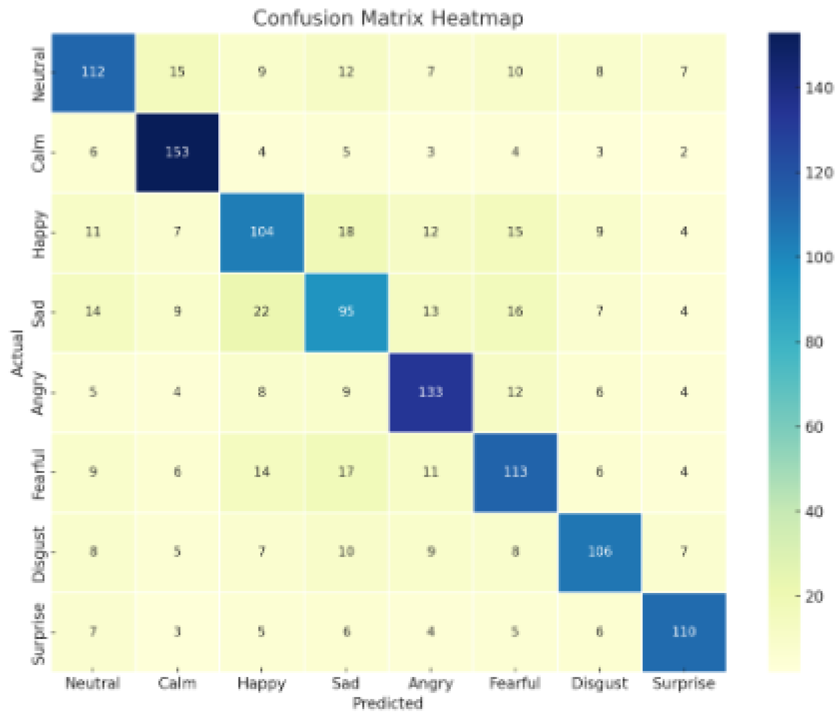


**Fig 4.11** Confusion Matrix for RAVDESS Dataset

### 4.2.3 Result Analysis for Combined Dataset (RAVDESS+SAVEE)

```
Test Accuracy: 60.70%
```

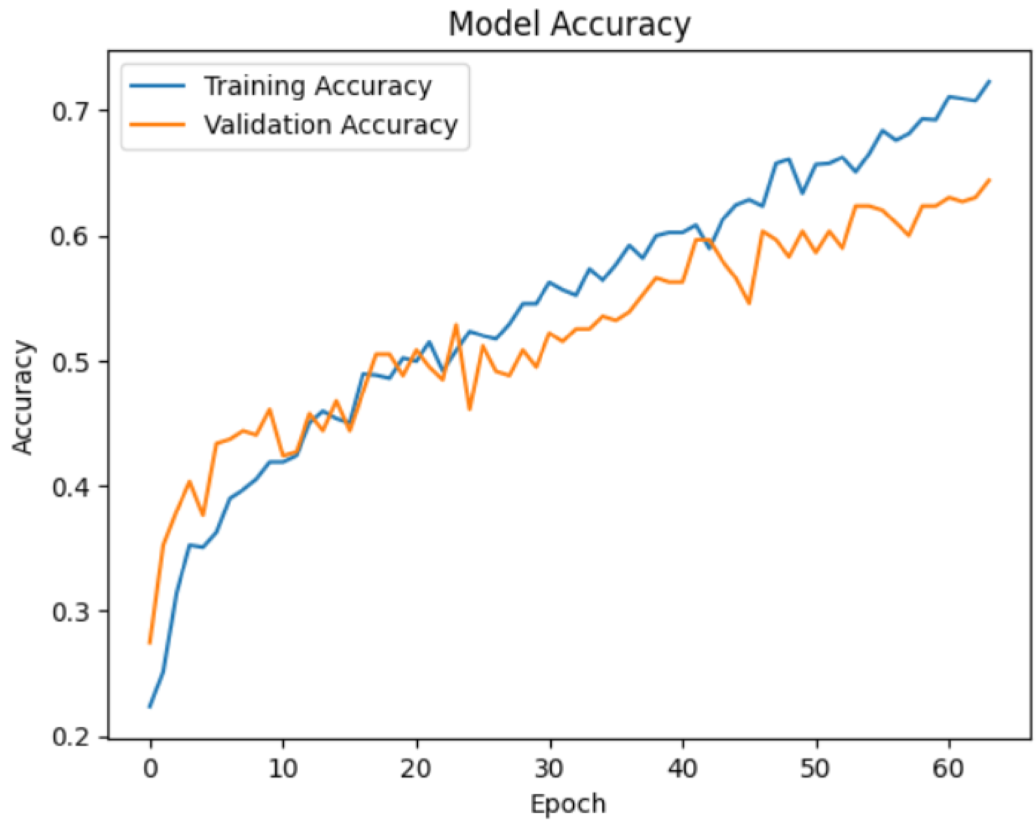


**Fig 4.11** CNN accuracy graph for Combined Dataset

```
Classification Report:
              precision    recall  f1-score   support

       angry       0.53      0.43      0.48        23
     disgust       0.76      0.62      0.68        63
     fearful       0.57      0.53      0.55        38
       happy       0.29      0.26      0.27        38
     neutral       0.75      0.86      0.80       130
         sad       0.24      0.21      0.22        39
    surprise       0.53      0.66      0.59        38

    accuracy                           0.61       369
   macro avg       0.52      0.51      0.51       369
weighted avg       0.60      0.61      0.60       369
```

**Fig 4.12** Classification Report for Combined Dataset

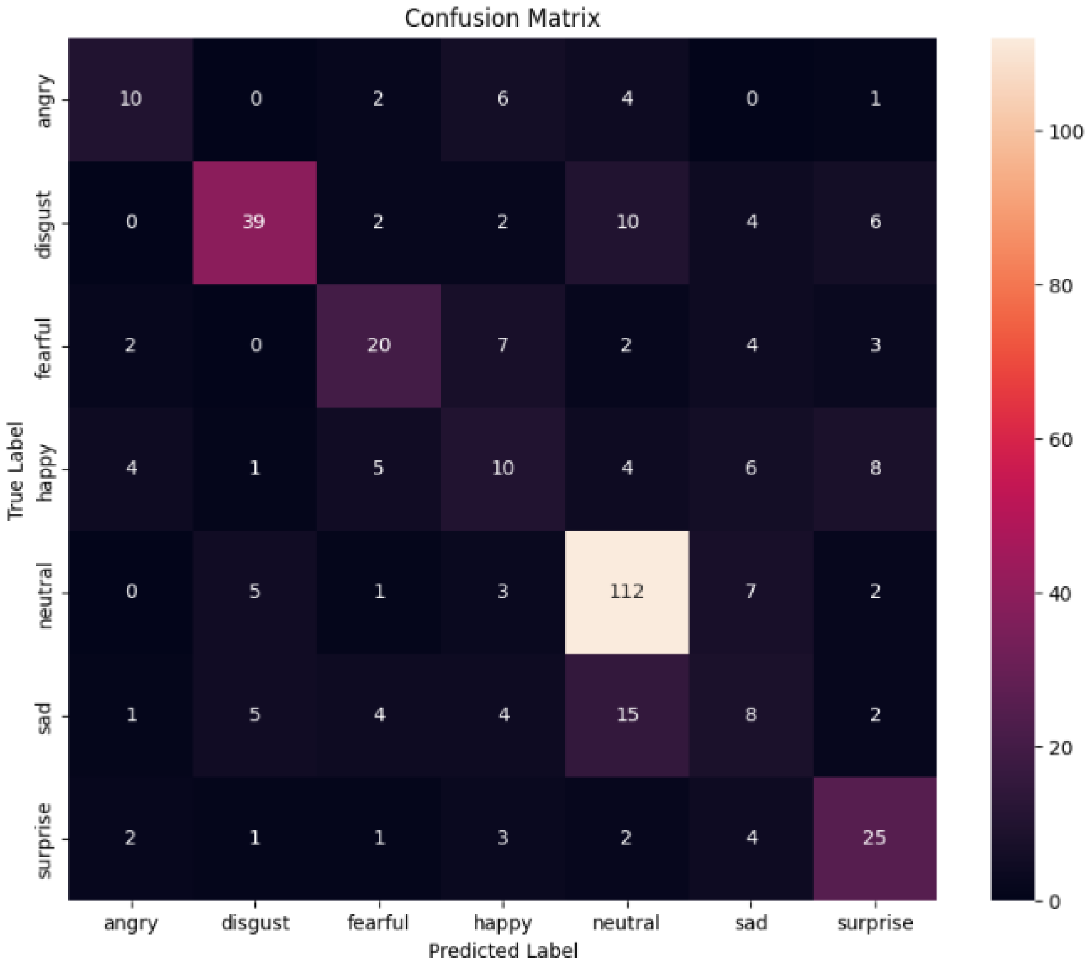


**Fig 4.13** Confusion Matrix for Combined Dataset

## Trained Model Prediction

In [75]:
```
finaldf[58:68]
```

Out[75]:

| | actualvalues | predictedvalues |
|---|---|---|
| 58 | male_fearful | male_happy |
| 59 | male_fearful | male_fearful |
| 60 | male_fearful | male_fearful |
| 61 | male_fearful | male_fearful |
| 62 | male_sad | male_sad |
| 63 | male_fearful | male_fearful |
| 64 | male_happy | male_happy |
| 65 | female_angry | female_angry |
| 66 | female_angry | female_fearful |
| 67 | male_angry | male_angry |

**Fig 4.14 Prediction**

# CHAPTER 5
# CONCLUSION AND FUTURE SCOPE

**Text Based-**

Sentiment analysis is the process of identifying and interpreting people's opinions, attitudes, and emotional expressions in text. These sentiments are generally classified as either positive or negative. Features such as parts of speech, particularly adjectives, are commonly used to detect sentiment in textual data. However, when adjectives are used alongside adverbs, determining the precise sentiment can become more challenging. In the proposed system, tweets are first extracted from a user's Twitter feed. The system then analyzes the frequency of each term within the tweet. Finally, a supervised machine learning method is used to perform the sentiment classification. **The average accuracy provided by our classifier is 99.8%.** When it comes to sentiment analysis of Twitter data, our built classifier is therefore completely trained and prepared.

**Speech Based -**

In this we demonstrated the application of Machine Learning in extracting emotions from speech audio data, shedding light on the manifestation of human emotions through voice.

The analysis highlighted the **CNN model's accuracy of around 97.92%** for **SAVEE Dataset, 66% accuracy for RAVDESS** Dataset and **60.70%** accuracy for combined **RAVDESS and SAVEE Dataset**. Notably, when categorizing data based on gender or specific emotions (like male angry, female happy), there was a marked increase in accuracy. With a remarkable **98% accuracy** rate, the model demonstrated remarkable performance in differentiating between audio from male and female speakers. Training with a wider variety of audio recordings may help improve accuracy even further.

This technology has possible applications in multiple domains. There are certainly a variety of applications for this technology in many domains. For instance, contact centres can utilize this technology for marketing or customer service improvement. They could leverage this technology in chatbots or voice-based virtual assistants to drastically increase their language comprehension and interaction capabilities. Moreover, without any copying problems, this technology could also be useful for linguistic research, explaining aspects of language comprehension and communication. Finally, possibly diversifying the dataset could improve the model's accuracy and utility further.

## FUTURE SCOPE

a. **Complex Text Pre-processing Methods:**
Examine more complex text preparation methods to enhance the quality of text representations, such as managing negations, lemmatization, and stemming.

b**. Feature Engineering:**
In order to extract more semantic information from the text, try experimenting with other feature extraction techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings.

c. **Model selection and tuning:**
Research other machine learning models, such as Random Forests, Support Vector Machines (SVM), or deep learning architectures, such as Transformer models for sentiment analysis or LSTM (Long Short-Term Memory), and evaluate how well they perform in comparison to the Naive Bayes classifier.
Adjust the hyperparameters to maximize the selected model's performance.

**d. Handling unbalanced data:**
Apply strategies such as under sampling, oversampling, or algorithms that effectively manage class imbalance to rectify any imbalance in the dataset.

**e. Exploring Additional Acoustic Features**:
Investigating other acoustic attributes like RAS MFCC, LPCC, PLP, or Harmonic Cepstrum beyond MFCC can potentially enhance the accuracy of speech emotion recognition (SER).

**f. Leveraging Lexical Features**:
Combining lexical and acoustic models in an ensemble approach could amplify accuracy, especially for contextually nuanced emotional expressions.

**g. Noise Reduction**:
Developing methods to filter out random silence and background noise from audio clips can enhance model accuracy and precision.

**h. Diversified Accent Training:**
Training models on a more diverse range of accents could broaden the model's applicability and effectiveness across various use cases.

# CHAPTER 6

# REFERENCES

[1] H. Zang, “The optimality of Naïve-Bayes”, Proc. FLAIRS, 2004

[2] C.D. Manning, P. Raghavan and H. Schütze, “Introduction to Information Retrieval”, Cambridge University Press, pp. 234-265, 2008

[3] A. McCallum and K. Nigam, “A comparison of event models for Naive Bayes text classification”, Proc. AAAI/ICML-98 Workshop on Learning for Text Categorization, pp. 41-48, 1998

[4] M. Schmidt, N. L. Roux and F. Bach, “Minimizing finite Sums with the Stochastic Average Gradient”, 2002

[5] Y. LeCun, L. Bottou, G. Orr and K. Muller, “Efficient BackProp”, Proc. In Neural Networks: Tricks of the trade 1998.

[6] T. Wu, C. Lin and R. Weng, “Probality estimates for multi-class classification by pairwise coupling”, Proc. JMLR-5, pp. 975-1005, 2004

[7] “Support Vector Machines” [Online], http://scikit learn.org/stable/modules/svm.html#svm-classification, Accessed Jan 2016

[8]P. Pang, L. Lee and S. Vaithyanathan, “Thumbs up? sentiment classification using machine learning techniques”, Proc. ACL-02 conference on Empirical methods in natural language processing, vol.10, pp. 79-86, 2002

[9] P. Pang and L. Lee, “Opinion Mining and Sentiment Analysis. Foundation and Trends in Information Retrieval”, vol. 2(1-2), pp.1-135, 2008

[10] E. Loper and S. Bird, “NLTK: the Natural Language Toolkit”, Proc. ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics ,vol. 1,pp. 63-70, 2002

[11] H. Wang, D. Can, F. Bar and S. Narayana, “A system for real-time Twitter sentiment analysis of 2012 U.S.presidential election cycle”, Proc. ACL 2012 System Demostration, pp. 115-120, 2012

[12] O. Almatrafi, S. Parack and B. Chavan, “Application of location-based sentiment analysis using Twitter for identifying trends towards Indian general elections 2014”. Proc. The 9th International Conference on Ubiquitous Information Management and Communication,2015

[13] L. Jiang, M. Yu, M. Zhou, X. Liu and T. Zhao, “Target-dependent twitter sentiment classification”, Proc. The 49th Annual Meeting of the Association

for Computational Linguistics: Human Language Technologies, vol. 1, pp. 151-160, 2011

[14] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou and P. Li, “User-level sentiment analysis incorporating social networks”, Proc. The 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1397- 1405, 2011

[15] A. Pak and P. Paroubek, “Twitter as a Corpus for Sentiment Analysis and Opinion Mining”, vol. 10, pp. 1320-1326, 2010

[16] B. Sun and TY. V. Ng, “Analyzing Sentimental influence of Posts on Social Networks”, Proc. The 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design, 2014

[17] Medha Khurana, Anurag Gulati & Saurabh Singh;” Sentiment Analysis Framework of Twitter Data Using Classification” 5th IEEE International Conference on Parallel, Distributed and Grid Computing(PDGC-2018), 20-22 Dec, 2018, Solan, India

[18] Satyendra Singh, Krishan Kumar & Brajesh Kumar;” Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques” 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), 26-27 May 2022.

[19] Prateek Garg; Sentiment Analysis of Twitter Data using NLTK in Python;

[20] Md. Rakibul Hasan1, Maisha Maliha2, M. Arifuzzaman3; Sentiment Analysis with NLP on Twitter Data, International Conference on Computer, Communication, Chemical,Materials and Electronic Engineering (IC4ME2), 11-12 July, 2019

[21] Anurag P. Jain & Mr. Vijay D. Katkar, Sentiments Analysis Of Twitter Data Using Data Mining, 2015 International Conference on Information Processing (ICIP)Vishwakarma Institute of Technology. Dec 16-19, 2015

[22] Sahar A. El_Rahman, Feddah Alhumaidi AlOtaibi & Wejdan Abdullah AlShehri, Sentiment Analysis of Twitter Data

[23] Huma Parveen & Prof. Shikha Pandey, Sentiment Analysis on Twitter Data-set using Naive Bayes Algorithm, 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)

[24] A. Barhan and A. Shakhomirov, “Methods for Sentiment Analysis of Twitter Messages”, Proc.12th Conference of FRUCT Association, 2012

[25] T. Mitchell, “Machine Learning”, McGraw Hill, 1997

[26] F. Jensen, “An Introduction to Bayesian Networks”, Springer, 1996

[27] T. C. Peng and C. C. Shih, “An Unsupervised Snippet-based Sentiment Classification Method for Chinese Unknown Phrases without using Reference Word Pairs”. IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology, vol. 3, pp. 243-248, 2010

[28] R. Feldman, “Techniques and applications for sentiment analysis”, Proc.

ACM, pp. 56-82, 2009

[29] https://github.com/christinechye/NLTK-sentiment-analysis/blob/main/NLTK_tutorial.ipynb

# Thesis_Final.pdf

Delhi Technological University

## Document Details

Submission ID

trn:oid:::27535:98025719

Submission Date

May 27, 2025, 9:46 PM GMT+5:30

Download Date

May 27, 2025, 10:08 PM GMT+5:30

File Name

Thesis_Final.pdf

File Size

1.8 MB

54 Pages

10,598 Words

63,945 Characters

# 5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

## Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 14 words)

## Exclusions

- 10 Excluded Matches

## Match Groups

**25** Not Cited or Quoted 5%
Matches with neither in-text citation nor quotation marks

**0** Missing Quotations 0%
Matches that are still very similar to source material

**0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation

**0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

## Top Sources

3% Internet sources
1% Publications
4% Submitted works (Student Papers)

## Integrity Flags

**1 Integrity Flag for Review**

**Hidden Text**
61 suspect characters on 2 pages
Text is altered to blend into the white background of the document.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

## Match Groups

- **25** Not Cited or Quoted 5%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

## Top Sources

| | |
|---|---|
| 3% | Internet sources |
| 1% | Publications |
| 4% | Submitted works (Student Papers) |

## Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

| | | | |
|---|---|---|---|
| 1 | Internet | tudr.thapar.edu:8080 | <1% |
| 2 | Internet | pdfs.semanticscholar.org | <1% |
| 3 | Submitted works | Jawaharlal Nehru University (JNU) on 2018-07-06 | <1% |
| 4 | Submitted works | Liverpool John Moores University on 2023-01-22 | <1% |
| 5 | Submitted works | Higher Education Commission Pakistan on 2022-09-30 | <1% |
| 6 | Internet | www.ijnrd.org | <1% |
| 7 | Submitted works | Liverpool John Moores University on 2022-10-26 | <1% |
| 8 | Publication | Dang, Yan, Yulei Zhang, and HsinChun Chen. "A Lexicon-Enhanced Method for Se... | <1% |
| 9 | Internet | ijirt.org | <1% |
| 10 | Internet | theses.liacs.nl | <1% |

11 Publication

Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelli... <1%

12 Submitted works

Universiti Teknologi MARA on 2020-07-10 <1%

13 Submitted works

University of North Texas on 2024-02-12 <1%

14 Internet

lirias.kuleuven.be <1%

15 Submitted works

Hellenic Mediterranean University on 2021-05-11 <1%

16 Internet

iieta.org <1%

17 Submitted works

Chandigarh University on 2018-06-22 <1%

18 Submitted works

University of Hertfordshire on 2022-10-02 <1%

19 Internet

insightsociety.org <1%

20 Internet

www.researchgate.net <1%

**DELHI TECHNOLOGICAL UNIVERSITY**

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42


## PLAGIARISM VERIFICATION

Title of the Thesis ____ "COMPUTATIONAL RECOGNITION OF EMOTIONS: SENTIMENT ANALYSIS IN HUMAN COMPUTER INTERACTION"

Total Pages: 53 Name of the Scholar : NAINA SHARMA Supervisor (s)

(1) Prof. Dinesh Kumar Vishwakarma

(2)____________________________________________

(3)____________________________________________

Department : INFORMATION TECHNOLOGY

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: Turnitin Similarity Index: 5 , Total Word Count: 10598

Date: 28.05.2025

**Candidate's Signature** **Signature of Supervisor(s)**

e-Receipt for State Bank Collect Payment

## REGISTRAR, DTU (RECEIPT A/C)

**BAWANA ROAD, SHAHABAD DAULATPUR, , DELHI-110042**
**Date: 28-May-2025**

| | |
|---|---|
| **SBCollect Reference Number :** | DUO1227890 |
| **Category :** | Miscellaneous Fees from students |
| **Amount :** | ₹3000 |
| **University Roll No :** | 2K23/ITY/06 |
| **Name of the student :** | NAINA SHARMA |
| **Academic Year :** | 2024-25 |
| **Branch Course :** | INFORMATION TECHNOLOGY |
| **Type/Name of fee :** | Others if any |
| **Remarks if any :** | M.TECH DISSERTATION FEES |
| **Mobile No. of the student :** | 7983719282 |
| **Fee Amount :** | 3000 |
| **Transaction charge :** | 0.00 |
| **Total Amount (In Figures) :** | 3,000.00 |
| **Total Amount (In words) :** | Rupees Three Thousand Only |
| **Remarks :** | M.TECH DISSERTATION FEES SUBMISSION 2025 |
| **Notification 1:** | Late Registration Fee, Hostel Room rent for internship, Hostel cooler rent, Transcript fee (Within 5 years Rs.1500/- & $150 in USD, More than 5 years but less than 10 years Rs.2500/- & $250 in USD, More than 10 |

than 10 years Rs.2500/- & $250 in USD, More than 10 years Rs.5000/- & $500 in USD) Additional copies Rs.200/- each & $20 in USD each, I-card fee,Character certificate Rs.500/-.

**Notification 2:**

Migration Certificate Rs.500/-, Bonafide certificate Rs.200/-, Special certificate (any other certificate not covered in above list) Rs.1000/-,Provisional certificate Rs.500/-, Duplicate Mark sheet (Within 5 years Rs.2500/- & $250 in USD, More than 5 years but less than 10 years Rs.4000/- & $400 in USD, More than 10 years Rs.10000/- & $1000 in USD)

**Thank you for choosing SB Collect. If you have any query / grievances regarding the transaction, please contact us**

**Toll-free helpline number i.e. 1800-1111-09 / 1800 - 1234/1800 2100**

**Email -:** **sbcollect@sbi.co.in**