SENTIMENT ANALYSIS OF SOCIAL MEDIA POSTS

A Mini Project Report Submitted

In partial fulfillment of the requirement for the award of the degree of

Bachelor of Technology In

Artificial Intelligence and Data Science by

DATHU PAVANSAI - 21N31A7217

GADADASU GANESH - 21N31A7221

GADDIPATI AKANKSHA - 21N31A7222

Under the Guidance of

Mr. CH. KANNAIAH

ASSISTANT PROFESSOR

Computational Intelligence Department

MRCET

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

MALLA REDDY COLLEGE OF ENGINEERING AND TECHNOLOGY

(Affiliated to JNTU, Hyderabad)

ACCREDITED by AICTE-NBA

Maisammaguda, Dhulapally post, Secunderabad-500014.

2021-2025

DECLARATION

I hereby declare that the project entitled **"Sentiment Analysis Of Social Media Post"** submitted to **Malla Reddy College of Engineering and Technology,** affiliated to **Jawaharlal Nehru Technological University Hyderabad (JNTUH)** for the award of the degree of **Bachelor of Technology** in **Artificial Intelligence and Data science** is a result of original research work done by me.

It is further declared that the project report or any part thereof has not been previously submitted to any University or Institute for the award of degree or diploma.

DATHU PAVANSAI - 21N31A7217
GADADASU GANESH - 21N31A7221
GADDIPATI AKANKSHA - 21N31A7222

Malla Reddy College of Engineering and Technology

(Autonomous Institution-UGC, Govt. of India)

(Affiliated to JNTUH, Hyderabad, Approved by AICTE, NBA&NAAC with 'A' Grade)Maisammaguda, Kompally, Dhulapally, Secunderabad – 500100
Website: www.mrcet.ac.in

CERTIFICATE

This is to certify that this is the bonafide record of the project titled "**Sentiment Analysis of Social Media Post**", submitted by **G. Ganesh(21N31A7221),Gaddipati Akanksha(21N31A7222), Dathu Pavan Sai(21N31A7217)** of B.Tech in the partial fulfillment of the requirements for the degree of Bachelor of Technology in Artificial Intelligence and , Dept. of CI during the year 2024-2025. The results embodied in this project report have not been submitted to any other university or institute for the award of any degree or diploma.

CH. KANNAIAH
Internal guide

Dr. D. Sujatha
Head of the department

External Examiner

ACKNOWLEDGMENT

We feel honored and privileged to place our warm salutation to our college **Malla Reddy College of Engineering and technology (UGC-Autonomous)**, our Director **Dr. V.S.K. Reddy** who gave us the opportunity to have experience in engineering and profound technical knowledge.

We are indebted to our Principal **Dr. S. Srinivasa Rao** for providing us with facilities to do our project and his constant encouragement and moral support which motivated us to move forward with the project.

We would like to express our gratitude to our Head of the Department **Dr. D. Sujatha** for encouragingus in every aspect of our system development and helping us realize our full potential.

We would like to thank our mini project coordinator **Dr. T.Shivaratna**, Associate Professor as well as our internal guide **CH. Kannaiah**, Associate Professor for their structured guidance and never-ending encouragement. We are extremely grateful for valuable suggestions and unflinching co-operation throughout mini project work.

We would also like to thank all supporting staff of department of CI and all other departments who have been helpful directly or indirectly in making our mini project a success.

We would like to thank our parents and friends who have helped us with their valuable suggestions and support has been very helpful in various phases of the completion of the mini project.

DATHU PAVANSAI - 21N31A7217

GADADASU GANESH - 21N31A7221

GADDIPATI AKANKSHA - 21N31A7222

ABSTRACT

Sentiment analysis of social media is a crucial process that utilizes natural language processing (NLP) and machine learning techniques to evaluate emotions and opinions expressed in user-generated content across platforms like Twitter, Facebook, and Instagram. This analysis categorizes sentiments into three primary classes: positive, neutral, and negative, providing insights into public perception of brands, products, or services.The methodology involves several steps: monitoring social media mentions, analyzing sentiment through advanced algorithms, and calculating a social sentiment score to quantify overall feelings. By extracting actionable insights from vast amounts of data, businesses can identify trends, enhance customer engagement, and improve marketing strategies.The significance of sentiment analysis lies in its ability to inform decision-making processes by revealing emotional drivers behind consumer behavior. As organizations increasingly rely on social media for brand management, effective sentiment analysis tools become essential for understanding audience perceptions and addressing potential issues proactively.In summary, sentiment analysis serves as a powerful tool for businesses to gauge public sentiment, track brand reputation, and refine their strategies based on real-time feedback from social media interactions.

TABLE OF CONTENTS

S. No.	Topic	Page No.
CHAPTER 1:	INTRODUCTION	1-5
	1.1 Purpose	2
	1.2 Background of project	2
	1.3 Scope of project	4
	1.4 Project features	5
CHAPTER 2:	SYSTEM REQUIREMENTS	6-9
	2.1 Hardware Requirements	5
	2.2 Software Requirements	5
	2.3 Existing System	5
	2.4 Proposed System	6
	2.5 Functional Requirements	7
	2.6 Non - Functional Requirements	8
CHAPTER 3:	SYSTEM DESIGN	9-12
	3.1 System Architecture	9
	3.2. UML Diagrams	10-12
CHAPTER 4:	IMPLEMENTATION	15-17
	4.1 Source Code	15-17
	4.2 Output Screens	19-22
CHAPTER 5:	CONCLUSION	23-24
	5.1 Conclusion	23
	5.2 Future Scope	24
	References	25

CHAPTER 1

1. INTRODUTION

Sentiment analysis of social media involves using natural language processing (NLP) and machine learning techniques to evaluate the emotions and opinions expressed in user-generated content across platforms like Twitter, Facebook, and Instagram. This analysis categorizes sentiments into three main classes: positive, neutral, and negative, providing valuable insights into public perception of brands, products, or services.The primary purpose of sentiment analysis is to understand how audiences feel about a brand or topic, which is crucial for businesses aiming to improve customer engagement and brand reputation. By analyzing social media interactions—such as likes, comments, shares, and mentions—organizations can gain actionable insights that inform marketing strategies and enhance customer service.The process typically involves monitoring social media mentions, collecting data through advanced algorithms, and calculating sentiment scores to quantify overall feelings. Tools designed for sentiment analysis automate these tasks, allowing businesses to efficiently sift through vast amounts of data and identify trends in public sentiment.In summary, sentiment analysis serves as a powerful tool for organizations to gauge public opinion, manage reputations, and make data-driven decisions based on real-time feedback from social media interactionsPURPOSE

The purpose of sentiment analysis of social media posts is to understand public emotions and opinions about brands, products, or services. Key objectives include:

- **Understanding Customer Perception**: Gaining insights into how offerings are viewed by customers.
- **Improving Reputation Management**: Proactively addressing negative feedback to maintain a positive online presence.
- **Enhancing Customer Service**: Identifying customer grievances and satisfaction levels for better service.
- **Market Research and Competitor Analysis**: Analyzing trends and competitor performance through public sentiment.
- **Informed Decision-Making**: Using emotional insights to make data-driven marketing and business decisions.

1.1 BACKGROUND OF PROJECT

Sentiment analysis of social media is a specialized area within natural language processing (NLP) that focuses on extracting subjective information from user-generated content across various platforms like Twitter, Facebook, and Instagram. This process involves identifying and categorizing sentiments expressed in posts as positive, negative, or neutral.The rise of social media has transformed how individuals express opinions and emotions, making these platforms rich sources of data for businesses and organizations. By utilizing machine learning and computational linguistics techniques, sentiment analysis enables the detection of emotional tones and public perceptions related to brands, products, or services.Organizations can collect social media data through APIs and apply sentiment analysis techniques to gain insights into customer feelings and behaviors. This understanding is crucial for tracking brand reputation, improving marketing strategies, and enhancing customer engagement. Research indicates that emotional factors significantly influence purchasing decisions, with approximately 70% of customer choices driven by emotions rather than rational considerations.As businesses increasingly recognize the importance of monitoring social sentiment, various tools have emerged to facilitate this analysis. These tools help organizations capture real-time feedback, identify trends, and respond proactively to customer sentiments. Overall, sentiment analysis serves as a vital component for brands aiming to connect with their audience and adapt to changing consumer preferences in an increasingly competitive digital landscape

1.3 SCOPE OF PROJECT

The scope of sentiment analysis in social media is extensive, encompassing various applications and benefits for businesses and organizations. Here are the key areas where sentiment analysis is impactful:

- **Understanding Customer Sentiment**: Sentiment analysis helps organizations gauge how their products, services, or brands are perceived by customers on social media. By analyzing posts, comments, and interactions, businesses can identify customer pain points and satisfaction levels.
- **Monitoring Brand Reputation**: This analysis allows companies to track public sentiment towards their brand in real-time. By categorizing mentions as positive, negative, or neutral, businesses can proactively manage their reputation and address any emerging issues.
- **Enhancing Marketing Strategies:** Insights gained from sentiment analysis can inform marketing campaigns by identifying what resonates with audiences. Brands can tailor their messaging and content strategies based on the emotional responses elicited from their target demographics.

- **Crisis Management**: Sentiment analysis enables organizations to detect potential PR crises early by monitoring negative sentiments. This allows for timely interventions to mitigate damage to brand reputation.
- **Competitor Analysis**: By analyzing sentiments related to competitors, businesses can gain insights into market positioning and consumer preferences, helping them adjust their strategies accordingly.
- **Product Development**: Feedback derived from sentiment analysis can guide product improvements and innovations by highlighting features that customers value or dislike.
- **Real-Time Feedback Loop**: Social media sentiment analysis provides a continuous stream of feedback that organizations can use to refine their strategies and enhance customer engagement dynamically.
- **Emotion Classification**: Beyond basic sentiment categorization (positive, negative, neutral), advanced sentiment analysis can classify emotions into more nuanced states (e.g., joy, anger, surprise), providing deeper insights into consumer feelings.
- **Data-Driven Decision Making**: By leveraging sentiment analysis, organizations can make informed decisions based on actual consumer sentiments rather than assumptions or outdated data

1.4 PROJECT FEATURES

- **Data Aggregation**: Sentiment analysis tools compile data from diverse social media platforms, such as Twitter, Facebook, and Instagram, thereby enriching the overall depth and breadth of the analysis.
- **Natural Language Processing (NLP)**: NLP algorithms meticulously analyze and interpret text by deconstructing it into its fundamental components, facilitating the identification of context and emotional nuances embedded within the language.
- **Sentiment Categorization**: The processed data is systematically classified into three primary sentiments—positive, negative, or neutral—based on the emotional tone articulated in the posts.
- **Emotion Detection**: Advanced sentiment analysis extends beyond basic categorization to discern specific emotional states (e.g., joy, anger) present within the text, offering a more nuanced understanding of sentiments.
- **Aspect-Based Analysis**: This feature scrutinizes the context surrounding particular words or phrases, yielding a detailed comprehension of sentiments related to specific facets of a product or service.
- **Data Visualization**: Insights are presented through sophisticated graphs and charts, enhancing the interpretability of data trends for businesses.

- **Real-Time Monitoring**: Many tools provide instantaneous sentiment analysis, empowering organizations to react promptly to emerging trends and potential issues.
- **Emoji and Emoticon Interpretation**: Advanced tools adeptly analyze emojis and emoticons to capture the emotional context conveyed in posts more accurately.
- **Customization and Flexibility**: Tools can be tailored to address specific business needs, enabling targeted sentiment analysis across various industries.
- **Human Oversight**: Some systems incorporate human verification to ensure precision and mitigate the risk of misinterpretation in sentiment assessment.

CHAPTER 2

2. SYSTEM REQUIREMNTS

2.1 HARDWARE REQUIREMENTS

- Processor: Multi-core processor (Intel i5/i7)
- RAM: Minimum 4 GB
- Hard Disk: Recommended 1 – 2 GB

2.2 SOFTWARE REQUIREMENTS

- Operating System: Windows, Linux, Mac
- Technology: Python
- IDE: Jupyter Notebook, PyCharm, or Visual Studio Code
- Libraries: Streamlit, NumPy, Pandas, Joblib, Scikit-learn, TensorFlow

2.3 EXISTING SYSTEM

Existing projects focused on sentiment analysis of social media encounter numerous drawbacks and challenges. One major issue is data quality, as social media content is frequently noisy, informal, and unstructured, featuring spelling mistakes, slang, emojis, and abbreviations that can impede the effectiveness of traditional natural language processing (NLP) techniques. Additionally, contextual ambiguity complicates sentiment classification; sentiments can be context-dependent, leading to misinterpretations—sarcasm or irony, for instance, may cause positive phrases to be read as negative. There are also semantic parsing challenges, as accurately identifying sentiment targets and modifiers in social media texts is complex. Many existing systems struggle with this ambiguity and require sophisticated models to navigate these nuances effectively. Furthermore, the high dimensionality and noise present in social media posts can complicate sentiment classification due to excessive features and irrelevant data, with significant portions of posts lacking clear sentiment indicators. Negation handling poses another challenge; phrases containing negation words must be interpreted correctly to grasp the true sentiment being conveyed.

2.4 PROPOSED SYSTEM

The proposed sentiment analysis system for social media posts introduces several significant enhancements over existing tools. First, it emphasizes focused data collection by exclusively gathering data from Twitter, which ensures the relevance and accuracy of the sentiment analysis. The system features a user-friendly interface designed for non-technical users, allowing marketers and brand managers to easily engage with sentiment analysis without requiring specialized technical skills. Additionally, it offers phrase-level sentiment analysis, enabling users to input specific phrases for more detailed insights—something many current tools lack. The system also supports real-time analysis and visualization, providing immediate data processing and graphical representations of sentiment trends, which facilitates quick comprehension of public sentiment. Furthermore, it employs enhanced analytical depth through advanced NLP techniques that account for nuances such as sarcasm and cultural context, thereby improving accuracy compared to simpler models. Lastly, the system automates the data collection process from Twitter, saving time and ensuring that insights are current. Overall, this system presents a more accessible, accurate, and insightful approach to sentiment analysis than traditional tools.

2.5 FUNCTIONAL REQUIREMENTS

The functional requirements for a sentiment analysis system targeting social media posts encompass several critical features:

- **Data Collection**: The system must be capable of aggregating data from various social media platforms, such as Twitter, Facebook, and Instagram, to ensure a comprehensive dataset for analysis.
- **Natural Language Processing (NLP)**: It should employ advanced NLP algorithms to process and interpret text, breaking it down into individual components while identifying context and emotions behind the words.
- **Sentiment Classification**: The system must classify the processed data into categories—positive, negative, or neutral—based on the emotional tone expressed in the posts.
- **Emotion Detection**: Beyond basic sentiment classification, the system should be able to identify specific emotional states (e.g., joy, anger) to provide deeper insights into user sentiments.
- **Aspect-Based Analysis**: The ability to examine the context surrounding specific words or phrases is essential for understanding sentiments related to particular aspects of products or services.
- **Data Visualization**: The system should present results through intuitive graphs and charts, making it easier for users to interpret insights and trends.
- **Real-Time Monitoring**: It must offer real-time analysis capabilities, allowing businesses to respond

promptly to emerging trends and issues in public sentiment.

- **User-Friendly Interface**: The interface should be designed for non-technical users, enabling marketers and brand managers to engage with sentiment analysis easily.
- **Customization Options**: Users should have the ability to customize the tool according to their specific business needs and industry requirements.
- **Automation of Data Collection:** The system should automate the data gathering process from social media platforms, ensuring that insights are current and relevant.
- **Negation Handling**: It must accurately interpret negations in text to understand the true sentiment being expressed.
- **Human Verification**: Incorporating human oversight can enhance accuracy by ensuring that the analysis correctly interprets complex sentiments.

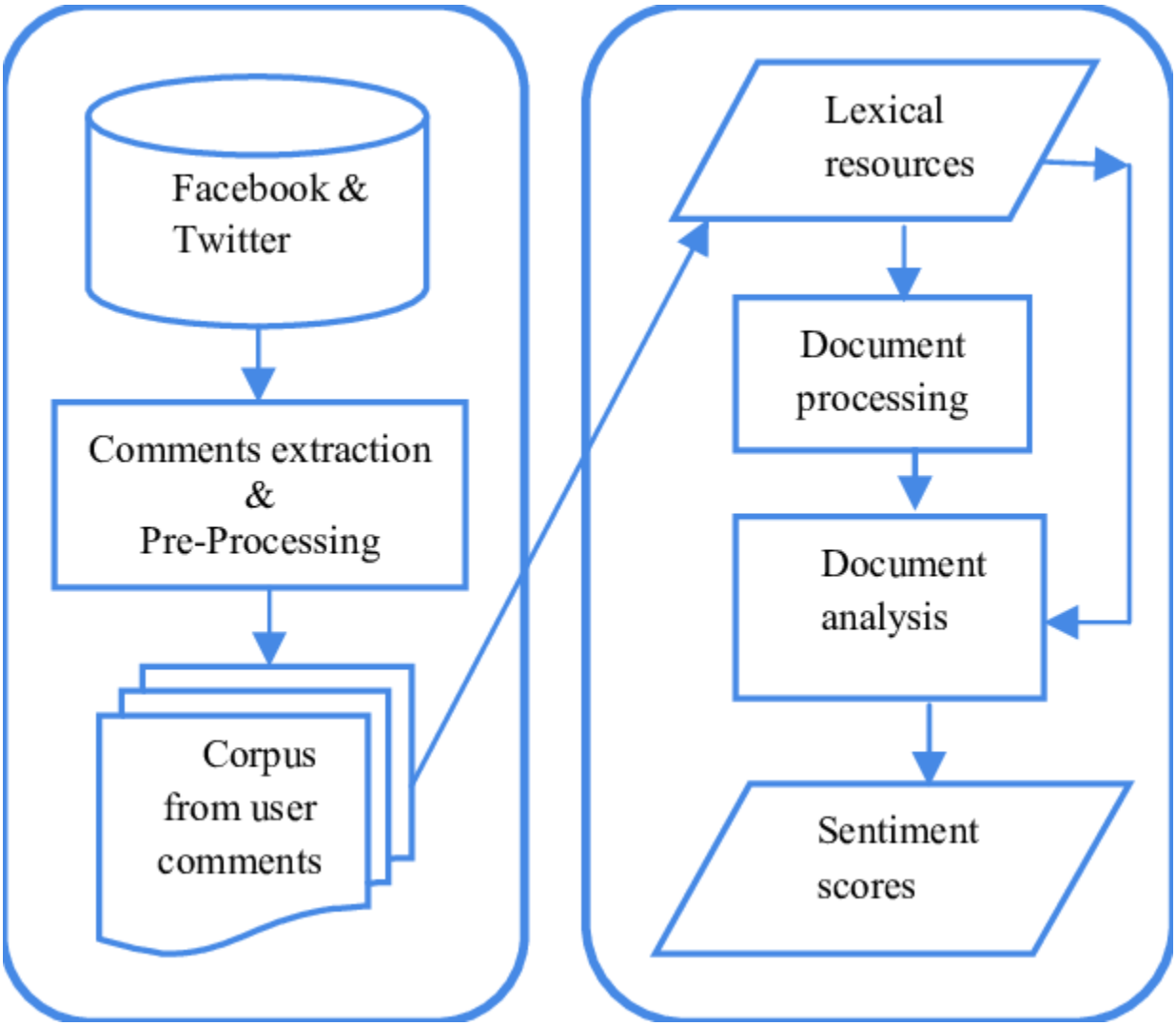
2.6 NON-FUNCTIONAL REQUIREMENTS

The non-functional requirements for a sentiment analysis system focused on social media posts encompass several critical aspects that ensure the system's overall performance, usability, and reliability.

- **Usability**: The system must be designed to be user-friendly, allowing individuals with varying levels of technical expertise to navigate and utilize the sentiment analysis tools effectively. This includes intuitive interfaces and clear instructions.
- **Performance**: The system should provide fast processing times for data collection and analysis, ensuring that users receive timely insights without significant delays, especially during peak usage periods.
- **Scalability**: It must be capable of handling increasing volumes of data as social media activity grows, allowing for the analysis of large datasets without degradation in performance.
- **Security**: The system must ensure the confidentiality and integrity of user data, implementing robust security measures to protect sensitive information shared on social media platforms.
- **Reliability**: The sentiment analysis tool should consistently deliver accurate results over time, minimizing errors and ensuring dependable performance in various conditions.
- **Maintainability**: The system should be designed for easy updates and maintenance, allowing developers to implement new features or fix issues without significant downtime.

CHAPTER 3
3. SYSTEM DESIGN

3.1 SYSTEM ARCHITECTURE

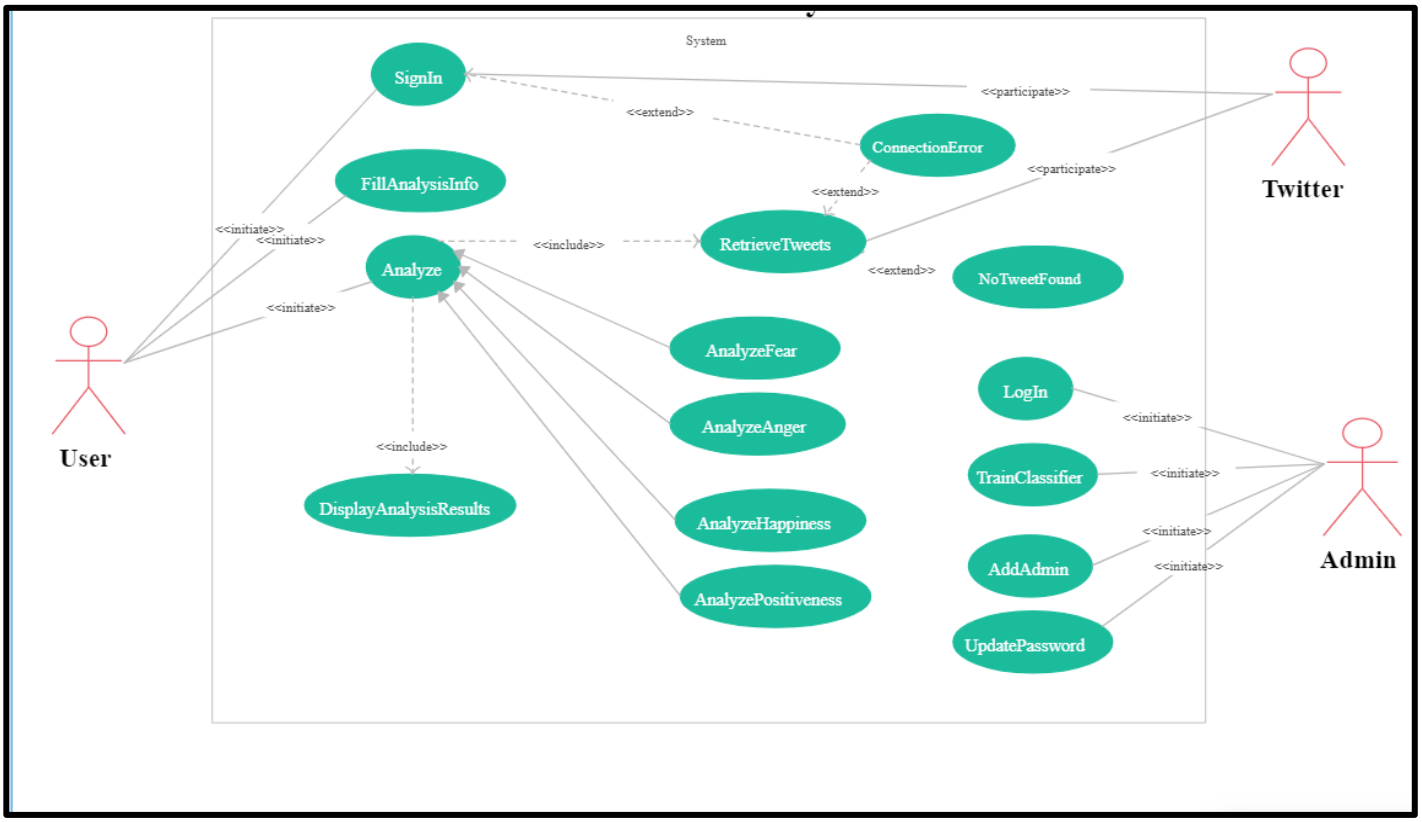


3.1: System Architecture

3.2 UML DIAGRAMS

3.2.1 USE CASE DIAGRAM

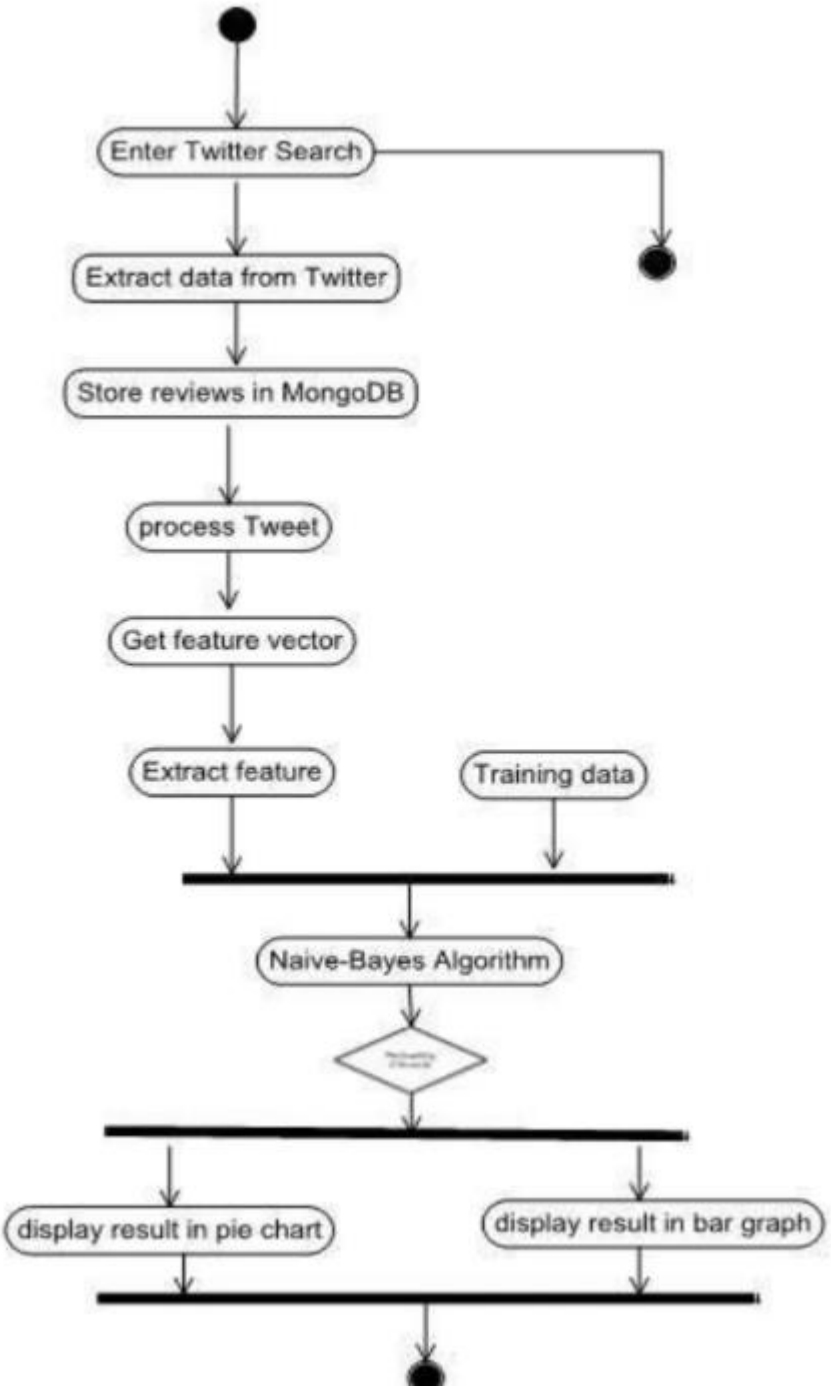
The use case diagram deals with the front-end working of the system. It depicts how system appears to a user. Here user is the actor that deals with the system. User will firstly login into the system after which she/he will browse tweets and select tweet for summarization.



3.2.1 : Use case Diagram

3.2.2 ACTIVITY DIAGRAM

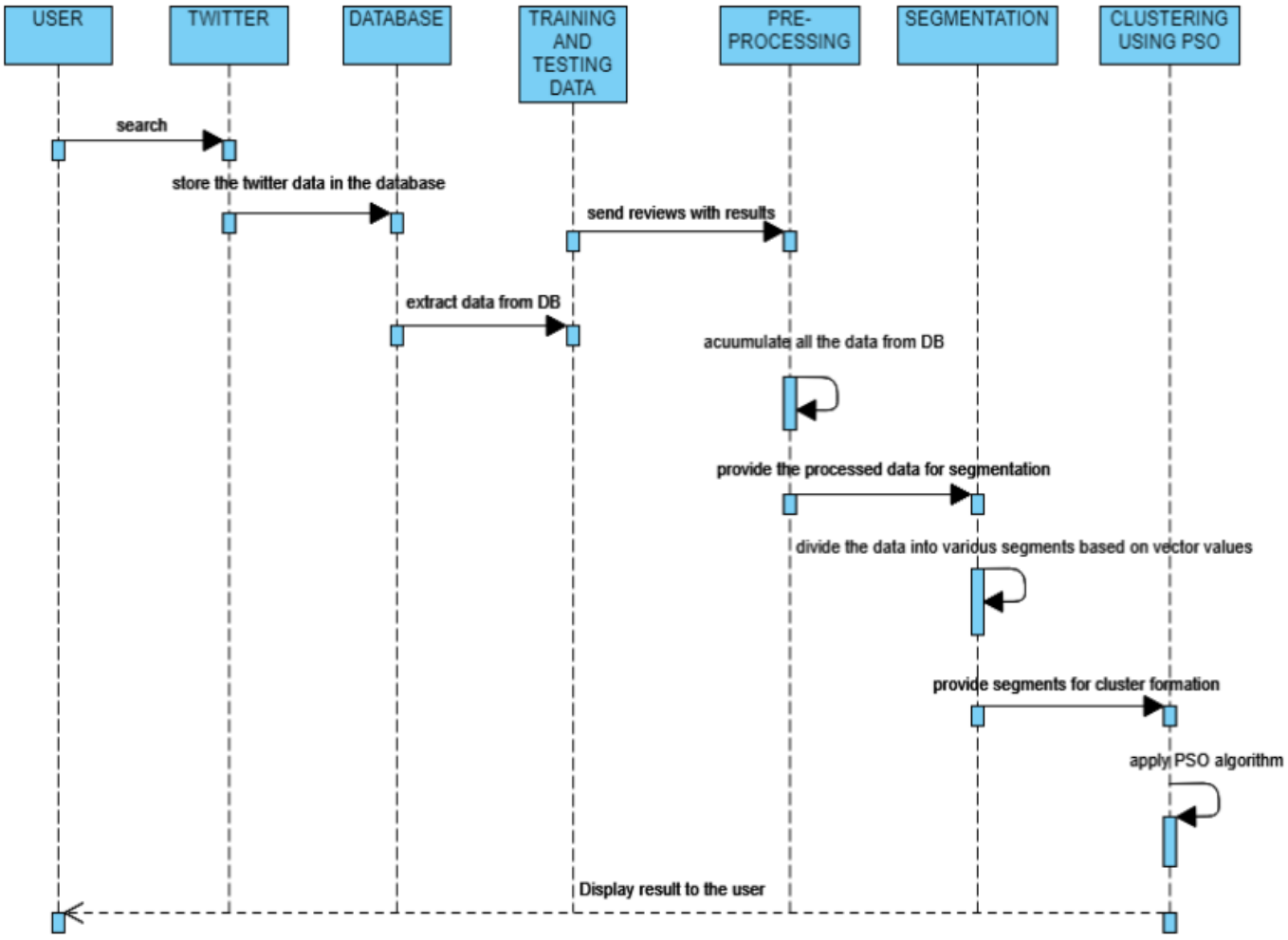
An activity diagram is a type of UML diagram that visually represents the flow of actions or activities within a system or a business process. Activity diagrams are particularly useful for modeling workflows, business processes, and the dynamic aspects of a system. They describe the sequence of activities and the flow of control between them.



3.2.2 : Activity Diagram

3.2.3 SEQUENCE DIAGRAM

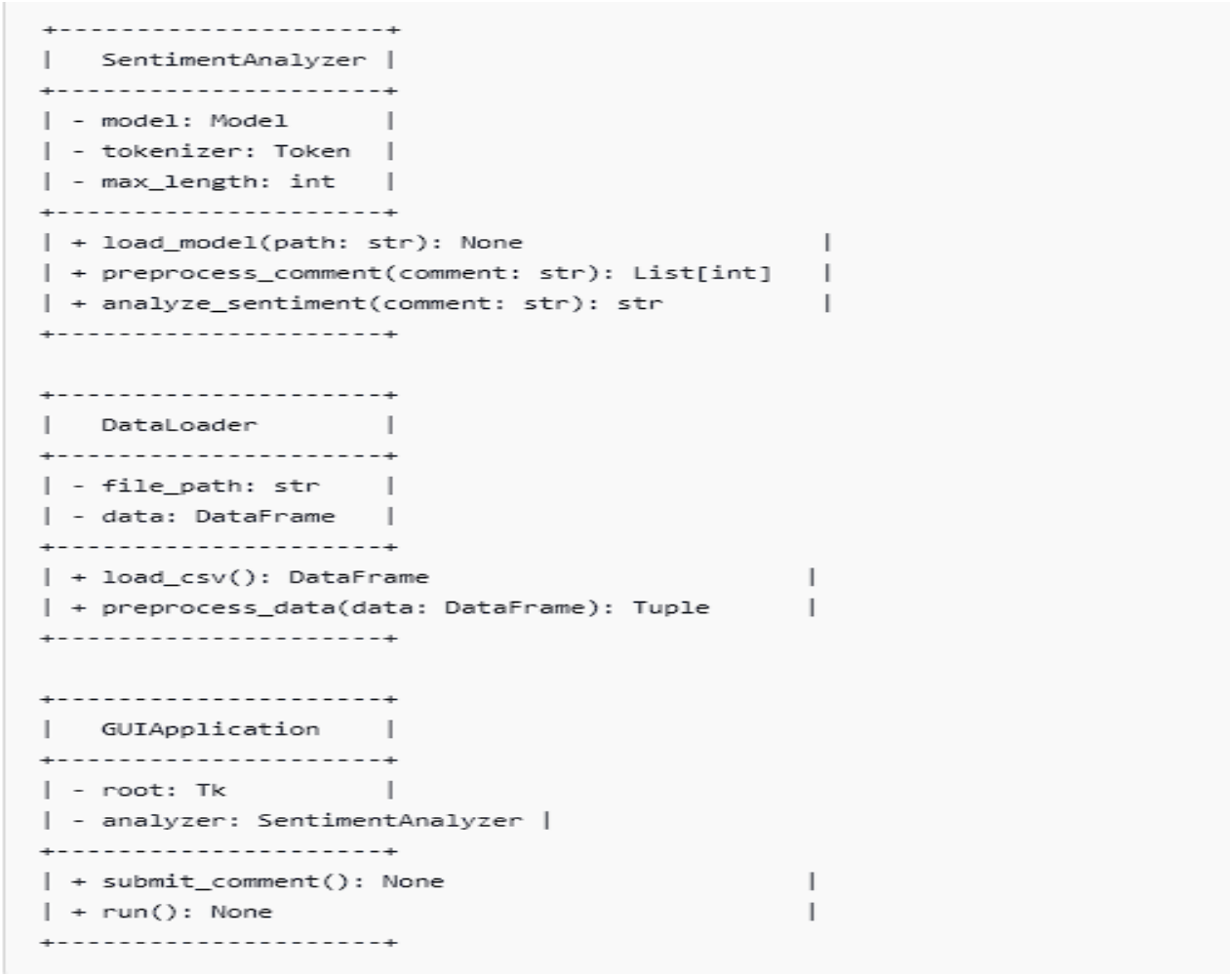
A sequence diagram is a Unified Modeling Language (UML) diagram that illustrates the sequence of messages between objects in an interaction. A sequence diagram consists of a group of objects that are represented by lifelines, and the messages that they exchange over time duringthe interaction. A sequence diagram shows the sequence of messages passed between objects. Sequence diagrams can also show the control structures between objects.



3.2.3 : Sequence Diagram

3.2.4 CLASS DIAGRAM

A class diagram is a type of static structure diagram in Unified Modeling Language (UML) that shows the structure of a system by visualizing its classes, their attributes, operations (or methods), and the relationships among the classes. Class diagrams are often used in software engineering to represent the blueprint of a program and show how different parts of the system interact with each other.



3.2.4 Class Diagram

CHAPTER-4

4. IMPLEMENTATION

4.1 CODE:

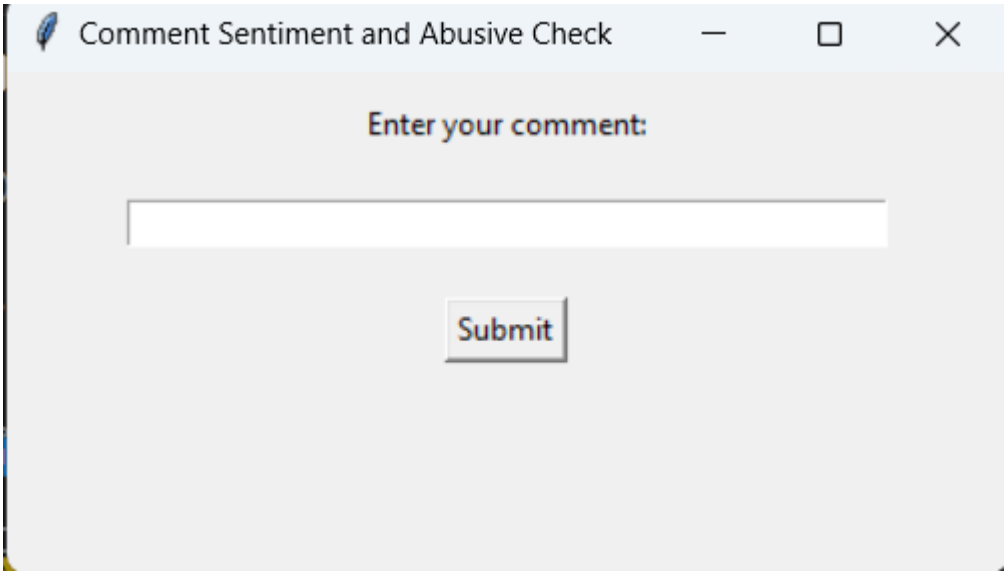
```
import pandas as pd
import tkinter as tk
from tkinter import messagebox
from nltk.sentiment import SentimentIntensityAnalyzer
from nltk.tokenize import word_tokenize
import nltk

# Download NLTK data files (only if not already downloaded)
nltk.download('vader_lexicon')
nltk.download('punkt')

# Initialize the VADER sentiment analyzer
sia = SentimentIntensityAnalyzer()

# Load abusive words from a CSV file
def load_abusive_words(file_path):
    try:
        df = pd.read_csv(file_path)
        # Assuming the abusive words are in a column named "words"
        abusive_words = set(df['words'].str.lower().dropna().tolist())
        return abusive_words
    except FileNotFoundError:
        messagebox.showerror("Error", "Abusive words file not found.")
```

```
        return set()
    except KeyError:
        messagebox.showerror("Error", "Column 'words' not found in the CSV file.")
        return set()

# Check if the comment contains any abusive language
def contains_abusive_language(comment, abusive_words):
    tokens = word_tokenize(comment.lower())
    for word in tokens:
        if word in abusive_words:
            return True
    return False

# Perform sentiment analysis using VADER
def analyze_sentiment(comment):
    sentiment_scores = sia.polarity_scores(comment)
    if sentiment_scores['compound'] >= 0.05:
        return "Positive comment"
    elif sentiment_scores['compound'] <= -0.05:
        return "Negative comment"
    else:
        return "Neutral comment"

# Handle comment submission in GUI
def submit_comment():
    comment = comment_entry.get()
    abusive_words = load_abusive_words("abusive_words.csv")  # Load abusive words file
```

```
    if contains_abusive_language(comment, abusive_words):
        messagebox.showerror("Abusive Comment", "The comment you typed contains abusive language and cannot be allowed.")
    else:
        vader_result = analyze_sentiment(comment)
        messagebox.showinfo("Sentiment Analysis Result", f"Comment sentiment: {vader_result}")

# Create the tkinter GUI window
root = tk.Tk()
root.title("Comment Sentiment and Abusive Check")
root.geometry("400x200")

# Label and entry for the comment
comment_label = tk.Label(root, text="Enter your comment:")
comment_label.pack(pady=10)

comment_entry = tk.Entry(root, width=50)
comment_entry.pack(pady=10)

# Submit button
submit_button = tk.Button(root, text="Submit", command=submit_comment)
submit_button.pack(pady=10)

# Run the GUI main loop
root.mainloop()
```

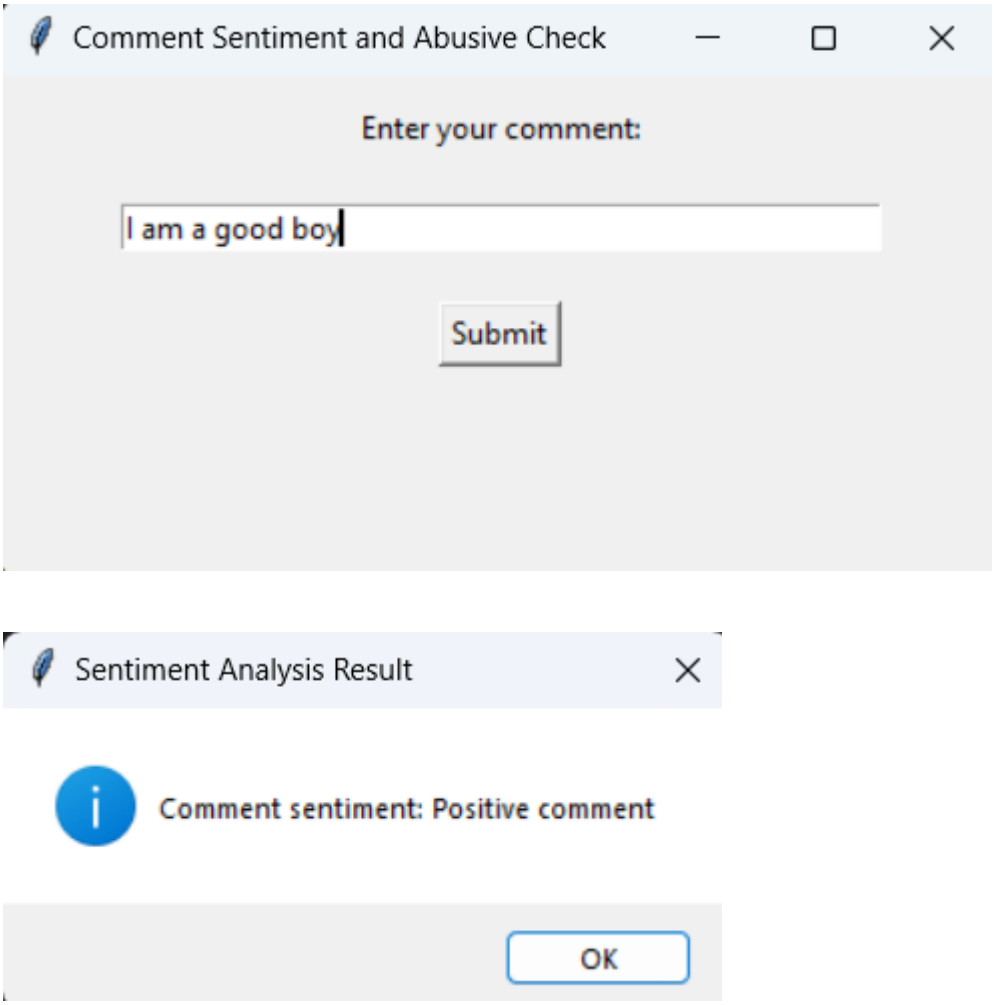
4.2 OUTPUT SCREENS

OUTPUT SCREEN 1:



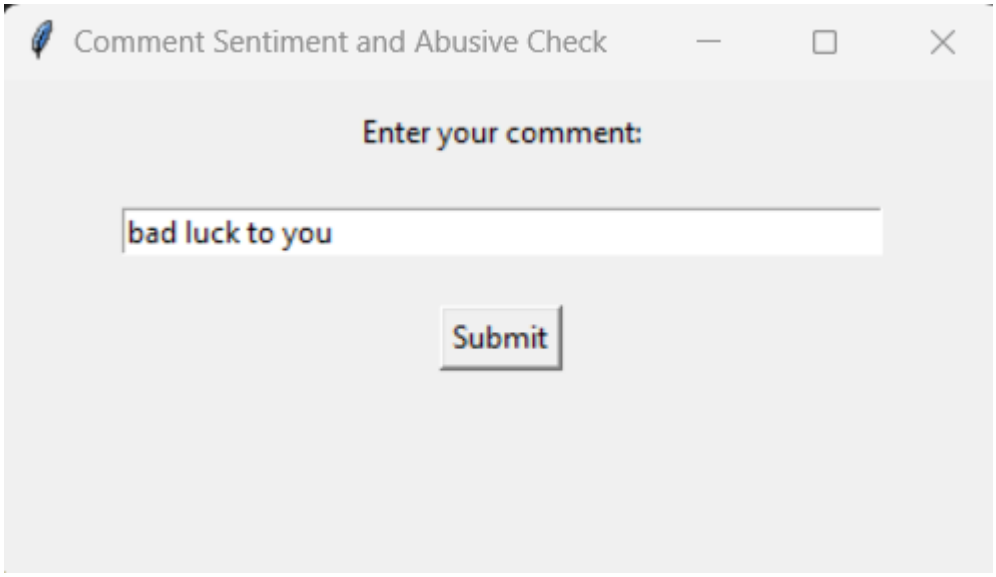
4.2.1 : Output screen1

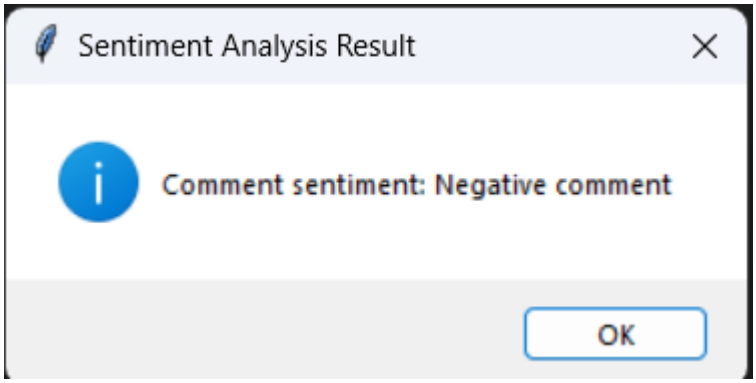
OUTPUT SCREEN 2:



4.2.2 : Output screen 2

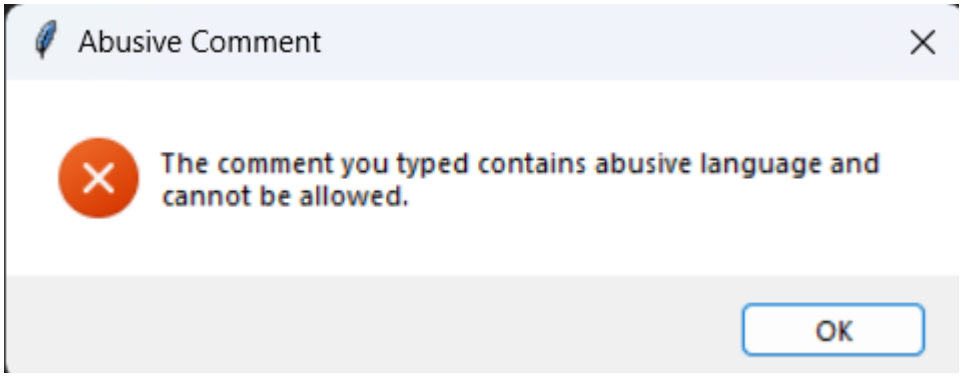
OUTPUT SCREEN 3:





4.2.3 : Output Screen 3

OUTPUT SCREEN 4:



4.2.4 : Output screen 4

CHAPTER 5

5.1 CONCLUSION

This project successfully implements a sentiment analysis tool with an integrated abusive language detector, designed to analyze comments entered by users. By leveraging the VADER sentiment analysis tool, the system efficiently classifies comments as positive, negative, or neutral based on sentiment scores. The integration of an abusive word filter, sourced from a CSV file, enhances the system's ability to flag and block inappropriate language, ensuring that only non-abusive comments proceed to the sentiment analysis stage.

The user-friendly GUI enables easy input and provides instant feedback on comment appropriateness and sentiment. This system holds practical applications in various domains, including social media management, customer feedback analysis, and online community moderation. Overall, the project demonstrates a robust and accessible approach to real-time sentiment and language filtering without the need for complex training datasets. This setup can be further extended by adding more features like advanced language filtering, dynamic sentiment scoring, or integration with databases to store flagged comments.

The project illustrates a streamlined yet effective solution to manage online interactions, promoting a more positive and respectful communication environment.

5.2 FUTURE SCOPE

The future scope of this sentiment analysis and abusive language detection project is expansive, with numerous opportunities for enhancement, particularly in content moderation, automated feedback systems, and real-time sentiment analysis applications. First, the abusive language detection could be significantly improved by incorporating real-time data sources or user-generated feedback, allowing the system to recognize evolving slang and context-specific terms that traditional keyword-based filters might miss. Implementing machine learning models that understand contextual abuse or implicit toxicity would make the detection even more sophisticated, allowing for a more nuanced and robust content filtering process.

Advances in natural language processing (NLP) offer promising enhancements to sentiment analysis, moving beyond VADER to state-of-the-art models such as BERT, RoBERTa, or GPT-based models. These models

can provide more accurate, context-aware sentiment analysis, adapting well to nuanced feedback typical of customer service or social media environments. Additionally, transfer learning with domain-specific datasets can further refine these models to meet the particular language and sentiment demands of social media and online communities, improving accuracy and relevance.

Expanding multilingual support is another essential area of growth. As social platforms and online communities are often multilingual, supporting sentiment and abusive language detection across multiple languages would make the tool more inclusive and applicable globally. This could be achieved by integrating multilingual models or utilizing translation tools that enable analysis beyond English, making it accessible to a broader user base.

Transforming the project into a web service or API would enhance its usability by allowing real-time integration with applications, websites, or social media platforms. This would enable businesses to leverage the tool as a background moderation service, automatically scanning user comments and providing immediate feedback on sentiment and language. Such an API could become a valuable component for companies looking to maintain a positive user environment and monitor their brand's online reputation.

User interface improvements could also make the tool more interactive and insightful. Features like comment history, flagged comment tracking, and data visualization can help community managers or moderators understand trends in user sentiment or types of detected abusive language. These visual insights can provide actionable data, aiding organizations in monitoring user feedback trends and responding accordingly.

REFERENCES

- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14). Association for the Advancement of Artificial Intelligence (AAAI).
- Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media, Inc
- Jurafsky, D., & Martin, J. H. (2021). Speech and Language Processing (3rd Edition). Draft chapters available at https://web.stanford.edu/~jurafsky/slp3/
- Fortuna, P., & Nunes, S. (2018). A Survey on Automatic Detection of Hate Speech in Text. ACM Computing Surveys (CSUR), 51(4), 1-30.

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of NAACL-HLT.
- Vidgen, B., & Derczynski, L. (2020). Directions in Abusive Language Training Data, a Systematic Review: Garbage In, Garbage Out. PLoS ONE, 15(12).