Sentiment Analysis of Social Media Presence

A PROJECT REPORT

Submitted by,

B Sai Kumar – 20211CBD0009
K Ganesh – 20211CBD0005
P Srinivas – 20211CBD0019
Y Sai Pallavi – 20211CBD0055

Under the guidance of,

Dr.Medikonda Swapna-Asso.Prof

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND TECHNOLOGY (BIG DATA)

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

MAY 2025

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Project report **"Sentiment Analysis of Social Media Presence"** being submitted by "Boyapati Sai Kumar, Koniki Ganesh, Pasupuleti Srinivas, Yarramsetty Sai Pallavi" bearing roll number(s) "20211CBD0009, 20211CBD0005, 20211CBD0019, 20211CBD0055" in partial fulfillment of the requiremnet for the award of the degree of Bachelor of Technology in Computer science and Technology (Big Data) is a bonafide work carried out under my supervision.

Dr. M Swapna
Asso. Prof - CSE
School of CSE&IS
Presidency University

Dr. S Pravinth Raja
Professor & HOD
School of CSE&IS
Presidency University

Dr. MYDHILI NAIR
Associate Dean
School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-VC School of Engineering
Dean -School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **Sentiment Analysis of Social Media Presence** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Science and Technology (Big Data)**, is a record of our own investigations carried under the guidance of **Dr. Swapna M, Asso. Prof, School of Computer Science Engineering & Information Science, Presidency University, Bengaluru.** We have not submitted the matter presented in this report anywhere for the award of any other Degree.

B Sai Kumar, 20211CBD0009 Sai Kumar

K Ganesh, 20211CBD0005 K. Ganesh

Y Sai Pallavi,20211CBD0055 Y. Saipallavi

P Srinivas, 20211CBD0019 Srinivas

ABSTRACT

The rise of social media platforms has led to an unprecedented volume of user-generated content, making sentiment analysis a crucial tool for understanding public opinion, brand perception, and social trends. This paper explores various machine learning and natural language processing (NLP) techniques used in sentiment analysis, including lexicon-based methods, supervised and unsupervised learning models, and hybrid approaches. A comparative analysis of existing sentiment analysis models, their accuracy, and application areas is provided. The study also discusses challenges such as handling sarcasm, multilingual data, and contextual ambiguity. The findings highlight the importance of sentiment analysis in fields like politics, healthcare, business intelligence, and crisis management.

Sentiment analysis of social media presence has become a critical research area, driven by the rapid growth of user-generated content on platforms like Twitter, Facebook, and Instagram. This study explores various sentiment analysis techniques, including lexicon-based, machine learning, and hybrid models, to classify social media posts into different sentiment categories. Sentiment classification accuracy can be improved with the help of artificial intelligence and deep learning approaches like Naïve Bayes, SVM, and LSTM.

However, despite these advancements, challenges remain in areas such as sarcasm detection, multilingual text processing, and real-time sentiment analysis. The research has underlined the importance of sentiment analysis in real-world applications such as marketing, politics, finance, healthcare, and crisis management. Future developments in explainable AI, cross-lingual analysis, and advanced deep learning techniques will further enhance the capabilities of sentiment analysis, making it an indispensable tool for businesses, researchers, and policymakers.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC, School of Engineering and Dean, School of Computer Science Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate **Dr. Mydhili Nair,** School of Computer Science Engineering & Information Science, Presidency University, and "**Dr. S Pravinth Raja**", Head of the Department, School of Computer Science Engineering & Information Science, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Dr. Swapna M, Asso. Prof,** and Reviewer **Ms. Arshiya Lubna**, School of Computer Science Engineering & Information Science, Presidency University for inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the CSE7301 Capstone Project Coordinators **Dr. Sampath A K, Mr. Md Zia Ur Rahman,** department Project Coordinators **"Ms. Suma N G"** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

B Sai Kumar

K Ganesh

Y Saipallavi

P Srinivas

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER- 1

INTRODUCTION

With the advent of the digital age, social media has also become an irreplaceable component of human life, which shapes opinions, forms perception, and drives conversations across the world. Twitter, Facebook, Instagram, and Reddit are just a few examples of such spaces where people post their thoughts, feelings, and responses towards different subjects of their interest, ranging from brands and products to politics and social matters. The sheer volume of user-created content on such sites offers a chance for businesses, organizations, and researchers to study public opinion and derive meaningful insights.This is a problem of creating a sentiment analysis solution tailored to analyzing the sentiment contained in the social media presence of individuals and organizations. Due to the massive influence of social media on individual and organizational reputation, learning the sentiment of social media posts, comments, and interactions has become vital to individuals and businesses. Sentiment analysis is the automatic detection of the sentiment or emotional tone expressed in text or speech. On social media, sentiment analysis can offer rich insights into public opinion, customer sentiments, and brand image. Through the analysis of the emotions expressed through social media postings, people and organizations can measure overall sentiment directions, recognize possible problems, and take relevant measures to preserve or improve their online reputation.

1.1 Significance of Sentiment Analysis

Sentiment analysis, or opinion mining, is a Natural Language Processing (NLP) methodology that consists of the computational analysis of individuals' opinions, feelings, and sentiments conveyed through textual data. Sentiment analysis categorizes sentiments as positive, negative, or neutral and aids in knowing how individuals feel about a particular subject or entity. This technology has been applied across various fields, including marketing, customer support, political punditry, and social awareness campaigns.

1.2 Objective of the Project

The main goal of this project, "Sentiment Analysis of Social Media Presence," is to create a strong sentiment analysis solution that can analyze social media content to provide valuable insights about public opinion. The project is centered on applying machine learning and NLP methods to process and analyze social media data efficiently. Through the determination of sentiment trends, companies can customize their strategies, policymakers can measure public sentiment, and organizations can improve customer interaction.

The research will discuss various methods of sentiment analysis, such as lexicon-based methods, machine learning algorithms, and deep learning approaches. In addition to that, it will also discuss some of the major challenges like preprocessing data, coping with sarcasm and unclear language, and coping with multilingual data. The result of the project will make an additional contribution to the evolving body of sentiment analysis by proposing a holistic methodology in analyzing social media sentiments as well as providing actionable suggestions to stakeholders.

In conclusion, this project highlights the importance of sentiment analysis in the current digital era and how it can revolutionize the manner in which institutions engage and react to public opinion. The conclusions obtained from this study will be the groundwork for continued innovation in the field, opening the way for more sophisticated and smarter sentiment analysis systems.

CHAPTER-2

LITERATURE SURVEY

To build up this model, we have read some earlier research papers.

Fuzzy Rule-based Unsupervised Sentiment Analysis from Social Media Posts, The paper [1]proposes a fuzzy logic-based unsupervised approach to sentiment classification, employing multiple lexicons and word sense disambiguation to classify posts as positive, negative, or neutral. While good with mixed datasets, its weaknesses include poor performance with short texts (tweets), no sarcasm detection, reliance on pre-defined lexicons, and poor support for multilingual data. For this, deep learning models like LSTM or BERT, hybrid lexicon-ML approaches, sarcasm detection approaches, and multilingual NLP approaches can be utilized to improve sentiment classification accuracy and contextual understanding.

Sentiment Analysis in Social Media and Its Application ,The paper [2] This is a systematic review of sentiment analysis methods, i.e., lexicon-based and opinion mining methods on Twitter data. It recognizes the extensive application of sentiment analysis in marketing, politics, and healthcare, but also recognizes the challenges of contextual ambiguity, detection of sarcasm, and the inability to adapt in real-time. To mitigate such challenges, deep learning models such as transformers (BERT, GPT), sentiment classification hybrid methods, and real-time streaming data processing can be used.

The paper[3] This article describes the application of AI to detect threats, disinformation, and hate speech on social media by machine learning and deep learning-based sentiment analysis. While AI enhances automated moderation and threat detection, challenges are bias in AI algorithms, privacy, and ethics in content filtering. For these, explainable AI (XAI), unbiased dataset selection, and ethical AI frameworks must be integrated to ensure fair and transparent sentiment analysis.

Impact of Social Media in Security and Crisis Management,
The paper[4] This paper explains how social media sentiment analysis can be used to aid crisis management by monitoring public emotions and reactions during crises. The research quotes the application of big data analytics and machine learning to derive social media insights but mentions challenges like the spread of misinformation, posting of fake news, and processing data in real-time. Solutions are fact-checking algorithms, real-time NLP models, and the application of geospatial analysis in combination with sentiment detection to improve crisis response plans.

A Systematic Review of Social Media-Based Sentiment Analysis, Emerging Trends and Challenges,The paper[5] This paper briefly discusses some of the techniques used in sentiment analysis, grouping them as lexicon-based, machine learning, and hybrid approaches and enumerating key issues in handling multilingual data, class imbalance, and real-time processing. It suggests that combining deep learning architectures with transfer learning (BERT, RoBERTa), data augmentation methods, and improved feature engineering can be used to enhance sentiment classification on diverse datasets.

A Review on Sentiment Analysis from Social Media Platforms,The paper [6] This paper briefly discusses some of the techniques used in sentiment analysis, grouping them as lexicon-based, machine learning, and hybrid approaches and enumerating key issues in handling multilingual data, class imbalance, and real-time processing. It suggests that combining deep learning architectures with transfer learning (BERT, RoBERTa), data augmentation methods, and improved feature engineering can be used to enhance sentiment classification on diverse datasets.

Beyond Positive or Negative: Qualitative Sentiment Analysis of Social Media Reactions to Unexpected Stressful Events,The paper [7] This work proposes a qualitative sentiment analysis with a contextual and affective interpretation rather than positive-negative classification alone. It proposes coping mechanism categories but is not real-time, automatic, or scalable. To these, hybrid qualitative-quantitative models, psychological NLP models, and automatic sentiment tagging systems can be employed to enhance sentiment comprehension.

A Model for Sentiment and Emotion Analysis of Unstructured Social Media Text,The paper [8] This book emphasizes machine learning and lexicon-based methods for sentiment and emotion extraction from social media. Naïve Bayes, SVM, and TF-IDF feature extraction are robust methods but are not context-sensitive, do not recognize sarcasm, and are not sensitive to complex sentence structures. Deep learning architectures such as CNNs, LSTMs, and attention-based transformers (BERT, GPT) can be used to increase context-sensitivity and sentiment accuracy.

Investigating Sentimental Relation Between Social Media Presence and Academic Success of Turkish Universities,

The paper[9] This study explores the connection between academic performance and social media sentiment using statistical sentiment analysis methods to analyze the reputation of universities. It lacks depth learning-based sentiment tracking, real-time fine-tuning, and contextual sentiment analysis. The following can be improved by incorporating AI-based sentiment prediction, multi-source data analysis, and longitudinal sentiment studies to provide more accurate academic insights.

Sentiment Analysis on Social Media, The paper [10] This work proposes a low-resource sentiment analysis approach based on basic NLP and machine learning techniques like Naïve Bayes and SVM for sentiment analysis on Twitter. While sufficient for basic polarity classification, it lacks state-of-the-art context awareness, sarcasm identification, and multilinguality. The integration of deep learning models, sentiment-aware embeddings like Word2Vec, FastText, and transformer-based sentiment models can be a game-changer in terms of performance and real-world applicability.

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

Sentiment analysis of social media data has gained significant attention due to its applications in various domains, including business intelligence, crisis management, and public opinion analysis. While numerous techniques—ranging from lexicon-based approaches to advanced deep learning models—have been developed, several key research gaps remain, limiting their effectiveness.

Sarcasm and Contextual Ambiguity: One of the biggest challenges in sentiment analysis is detecting sarcasm and handling contextual ambiguity. Existing machine learning and lexicon-based approaches struggle to differentiate between literal and sarcastic statements, often leading to misclassification of sentiments.

Multilingual and Code-Mixed Text Processing: Most sentiment analysis models are trained on English datasets and do not generalize well to other languages, especially in code-mixed scenarios where multiple languages are used in a single text. The lack of large, annotated datasets for underrepresented languages further exacerbates this issue.

Real-Time Sentiment Analysis: While deep learning models like LSTM, BERT, and GPT perform well in offline settings, their computational complexity makes real-time processing of large-scale social media data challenging. Efficient real-time sentiment analysis models that balance speed and accuracy are still an open research problem.

Bias and Ethical Concerns in AI Models: Many sentiment analysis models exhibit biases due to the datasets they are trained on. These biases can lead to unfair or inaccurate sentiment classifications, particularly for sensitive topics such as political discourse or social issues. Explainable AI (XAI) and fairness-aware training methodologies need further development.

Sentiment Analysis Beyond Polarity Classification: Traditional sentiment analysis models primarily focus on classifying text as positive, negative, or neutral. However, emotions are more complex and nuanced. Existing methods do not sufficiently capture emotions such as fear, anger, or joy, which are crucial for a deeper understanding of user sentiment.

Misinformation Detection and Crisis Management: Sentiment analysis is increasingly used in crisis management to track public opinion and misinformation. However, current models lack mechanisms to filter out fake news, manipulated content, or bot-generated sentiments, which can skew the analysis results.

Lack of Adaptive Learning Mechanisms:The prevailing sentiment analysis models tend to use static databases, hence their relative inability to cope with dynamically changing language habits, slang usage, and the arising social themes. Their restricted capacity to constantly learn from ongoing new data inputs without retraining restricts them to long-term efficacy and correctness. Creating autonomous learning models able to adapt continuously to real-world shifts in expressions of sentiment remains a key focus area.

Integration of Multimodal Data: Most sentiment analysis methodologies are concentrated only on text-based data without any consideration for other forms of sentiment expression through media like images, videos, or audio. Most social media users also share views through voice notes, emojis, GIFs, and memes, so integrating multimodal analysis would be very necessary to give a better overview of sentiment.

Scalability Challenges with Big Data Datasets: As social media data increasingly grows with time, efficient handling of huge datasets continues to be a challenge. Most sentiment analysis methods used today do not scale when, for example, used with real-world data streams consisting of millions of posts per second. Studies in distributed computing, cloud-based sentiment analysis, and efficient deep learning architectures can ensure scalability issues are dealt with.

CHAPTER-4

PROPOSED MOTHODOLOGY

This research aims to develop an efficient sentiment analysis system for social media data using machine learning and natural language processing (NLP) techniques. The methodology consists of multiple phases, including data collection, preprocessing, model training, and deployment.

1. Data Collection

- **Data Sources**: Twitter (via Tweepy API), social media websites (via Scrapy, if applicable).
- **Dataset**: Pre-collected datasets from Kaggle will be used for training and validation.
- **Data Format**: Social media text, including tweets and posts, with sentiment labels (positive, negative, neutral).

2. Data Preprocessing

Before applying machine learning models, raw text data undergoes preprocessing to improve accuracy:

- **Tokenization**: Splitting text into individual words.
- **Stopword Removal**: Eliminating common words (e.g., "is," "the") using NLTK.
- **Lemmatization/Stemming**: Converting words to their base forms.
- **Handling Emojis and Special Characters**: Replacing emojis with their sentiment labels and removing unnecessary symbols.
- **Removing URLs and Mentions**: Eliminating links and user mentions to retain only relevant content.

3. Feature Extraction

- **TF-IDF (Term Frequency-Inverse Document Frequency)**: To quantify text importance.
- **Word Embeddings (Optional)**: Pre-trained embeddings like Word2Vec or FastText can be used for improved representation.

4. Sentiment Classification Model

Two machine learning models will be implemented and compared:

- **Random Forest Classification**: A robust ensemble learning method that builds multiple decision trees for sentiment classification.
- **Logistic Regression**: A statistical method effective for binary classification tasks like positive vs. negative sentiment.

5. Model Training and Evaluation

- **Training Data**: The collected dataset will be split into training and testing sets (e.g., 80% training, 20% testing).
- **Evaluation Metrics**: Accuracy, Precision, Recall, and F1-score will be used to measure model performance.

6. Backend and Frontend Development

- **Backend**: Python-based Flask/Django API to serve predictions.
- **Frontend**: A minimal web interface for users to input text and receive sentiment classification results.

7. Deployment and Real-Time Sentiment Analysis

- **Integration with Twitter API**: The trained model will be deployed to analyze live tweets.
- **Real-Time Processing**: The system will continuously fetch, analyze, and display sentiment trends.
- **Cloud Deployment (Optional)**: The final application may be hosted on cloud platforms like AWS or Heroku.

By implementing this methodology, the project will provide an efficient and scalable sentiment analysis system that can be used for social media analytics, opinion mining, and trend prediction.

CHAPTER-5

OBJECTIVES

1. Develop a Robust Sentiment Analysis Model
    - Implement machine learning-based sentiment classification using Random Forest and Logistic Regression.
    - Utilize NLP techniques for feature extraction and text preprocessing to improve classification accuracy.
    - Optimize model performance using evaluation metrics like accuracy, precision, recall, and F1-score.

2. Handle Language and Contextual Challenges
    - Address challenges related to sarcasm, contextual ambiguity, and slang commonly found in social media text.
    - Implement preprocessing techniques such as stopword removal, lemmatization, and emoji handling.
    - Explore word embeddings or hybrid models for better contextual understanding.

3. Perform Real-Time Sentiment Tracking
    - Integrate with the Twitter API to fetch and analyze live tweets.
    - Develop a real-time sentiment monitoring system that processes and classifies social media data dynamically.
    - Ensure efficient handling of high-volume streaming data with optimized computational techniques.

4. Provide Visualization and Reporting Tools
    - Design a web-based interface to display sentiment analysis results interactively.
    - Implement graphical representations such as sentiment trend graphs and word clouds.
    - Enable users to track sentiment changes over time for data-driven decision-making.

5. Improve Sentiment Classification with Deep Learning
   - Build sophisticated deep learning models, including LSTMs and transformers, to enhance sentiment classification accuracy.
   - Fine-tune pre-trained models such as BERT and RoBERTa to improve comprehension of social media language.
   - Compare against performance of traditional machine learning approaches to establish optimal methods.

6. Enhance Multimodal Sentiment Analysis
   - Integrate image, video, and audio data with text to develop an integrated sentiment analysis system.
   - Leverage computer vision methods for analyzing visual sentiment in memes and GIFs.
   - Create models able to process and understand multimodal content across social media platforms.

7. Detect and Mitigate Bias in Sentiment Analysis
   - Examine biases in current sentiment analysis datasets and models.
   - Employ fairness-aware training practices to minimize biased sentiment labels.
   - Create explainable AI methods to enhance sentiment prediction transparency.

8. Implement Domain-Specific Sentiment Analysis
   - Adapt sentiment analysis models for industry-specific applications like healthcare, finance, and e-commerce.
   - Construct domain-adaptive lexicons and embeddings to strengthen sentiment classification for specialized domains.
   - Evaluate sentiment patterns in industry-specific conversations to draw useful insights.

Through these developments, this project seeks to develop more advanced methodologies for sentiment analysis, overcome challenges faced, and improve the understanding of social media sentiment as well as make it more accurate and complete.

Facebook usage during crises and emergencies

Facebook usage is very different from Twitter. In the majority of instances, information posted is accessible only to the community of the provider (his/her 'friends'). Information posted is more personal and has the tendency to open a discourse among individuals concerned or interested.

During crises and emergencies, it can be a significant communication channel with the advent of pages pertaining to some significant or stressful incidents, which enables communities to exchange support and resources when confronted to dramatic situations like storms or floods for example.

Facebook can also be utilized 'in time of peace' to offer training and information on prevention. This is the task of 'Community Managers', who take care of the accounts of public authorities and private companies with a clearly defined strategy.

Data Cleaning and Filtering

Throughout the data extraction phase, social media messages underwent a strict filtering process to incorporate only certain types of tweets, as suggested by Burgess and Bruns (2012). This categorization splits tweets into six categories:

1. Original tweets – posts that are not replies or retweets.

2. Retweets – messages containing "RT @user…" or similar markers.

3. Unedited retweets – retweets that start particularly with "RT @user…".

4. Edited retweets – retweets which have been edited and do not start with "RT @user…".

5. Actual replies – tweets that address another user with "@user" but are not retweets.

6. URL shares – tweets that contain web links.

For the purpose of this study, the analysis was restricted to original tweets, edited retweets, and actual replies, because these types generally have sufficient user-generated content to offer clues about individual experience or affective expression. Retweets that merely replicated content (unedited retweets, retweets), were authored by non-person accounts (i.e., news organizations), or copied verbatim several times per day were filtered out.

Additionally, material not from Twitter was also deleted. This data cleansing reduced the original dataset of 14,231 messages to a final 2,099 tweets, which was used as the final dataset for qualitative coding and analysis. This refined selection ensured a focus on unique, user-generated expressions during the crisis period.

Translation and Validation Process

Because the dataset contained tweets in German, a professional translator translated the content to provide linguistic accuracy and cultural appropriateness. To further confirm the maintenance of symbolic and cultural meanings, a second native German speaker who was also proficient in English reviewed the translations independently.

To further ensure the validity of the translated data, the coders themselves, who were also both English and German proficient, took an extra validation step. They randomly selected tweets written in German, translated them into English through automated translation tools, and compared the translations to the human-translated ones. Only after such intensive verification were the translated tweets incorporated into the next coding phase.

CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

1. System Design

The system is designed as a web-based sentiment analysis platform that processes social media data to determine sentiment polarity (positive, negative, or neutral). The architecture follows a modular approach, ensuring scalability and efficiency.

1.1 System Components

1. User Interface (Frontend)
   - A web-based interface where users can input text or analyze social media sentiment.
   - Displays sentiment analysis results with visualizations (charts, word clouds).
2. Backend (Flask/Django API)
   - Handles user requests and processes sentiment analysis.
   - Connects with the machine learning model and data sources.
3. Data Collection Module
   - Tweepy: Fetches tweets for real-time sentiment tracking.
   - Scrapy: (Optional) Scrapes social media content from websites.
   - Kaggle Datasets: Used for model training and evaluation.
4. Preprocessing & Feature Extraction
   - Uses NLTK for text cleaning, tokenization, and stopword removal.
   - TF-IDF and Word Embeddings help in converting text into numerical features.
5. Machine Learning Models
   - Random Forest Classification and Logistic Regression for sentiment classification.
   - Evaluated using accuracy, precision, recall, and F1-score.
6. Database (Optional)
   - Stores processed sentiment data for historical analysis.
7. Visualization & Reporting
   - Provides sentiment trend analysis using graphs and reports.

2. Implementation

The implementation follows an iterative development approach, ensuring continuous improvements.

2.1 Technology Stack

- Programming Language: Python (for ML and backend processing)
- Frameworks & Libraries:
    - NLP & ML: NLTK, Scikit-learn, TensorFlow/Keras (if needed)
    - Data Handling: Pandas, NumPy
    - Web Framework: Flask/Django for backend API
    - Frontend: HTML, CSS, JavaScript (React.js or simple UI)
    - Data Collection: Tweepy (Twitter API), Scrapy (Web Scraping)

2.2 Development Phases

1. Data Collection & Preprocessing
    - Fetch social media data using APIs (Twitter, Scrapy).
    - Clean text (removing stopwords, URLs, special characters).
    - Convert text into feature vectors using TF-IDF/Word Embeddings.
2. Model Training & Evaluation
    - Train Random Forest and Logistic Regression on Kaggle datasets.
    - Evaluate models and optimize hyperparameters.
3. API Development & Integration
    - Develop Flask/Django API to serve model predictions.
    - Integrate frontend for user interactions.
4. Visualization & Deployment
    - Create real-time dashboards with sentiment trends.
    - Deploy the system on AWS/Heroku for accessibility.

3. Expected Outcomes

- A real-time sentiment analysis system that can process social media data dynamically.
- Improved accuracy using hybrid NLP and machine learning models.
- User-friendly frontend with sentiment visualization.
- Potential future integration with explainable AI (XAI) and deep learning models.

Key Uses of AI in Mitigating Safety and Security on Social Media

Artificial Intelligence performs a pivotal role in improving social media safety and security using some advanced methods. One of its fundamental uses is automated content filtering, where the AI system detects and deletes violent content like hate speech, fake news, and violent images. These systems have the ability to analyze text, images, and videos to keep the online community safer and respectful (Imran et al., 2020).

Another principal use is for the prevention and detection of cyberbullying. AI applications track users' engagement and look for patterns in behavior that indicate harassment or abuse, allowing for early intervention (Benabdelouahed & Dakouan, 2020).

Identity verification is another important area where AI is valuable. Algorithms analyze user information to verify identities and identify suspicious or fictitious profiles, minimizing the opportunity for impersonation and identity fraud (Al-Ghamdi, 2021).

Furthermore, AI-driven anomaly detection tools assist in the identification of suspicious or potentially malicious activities like account breaches or bot automation. These systems improve the overall security infrastructure of social media platforms by allowing faster action on potential threats (Malik, 2020).

AI Methodologies and Techniques for Online Threat Management

Artificial intelligence applies an array of sophisticated techniques and tools to detect and deal with online threats efficiently. The most common method applied is machine learning, where an algorithm learns from historical data to recognize patterns and activities characteristic of likely security concerns (Perakakis & Mastorakis, 2019). Supervised models of learning, including those for classification and regression, are particularly useful to learn from tagged sets of data in order to properly categorize future data (Jorge & Ross, 2019).

Natural Language Processing (NLP) is also essential in these solutions since it enables systems to understand and process vast amounts of textual data. This feature facilitates the identification of toxic behavior, harmful language, and even the sentiment of user-generated content (Imran et al., 2020).

In addition, anomaly detection methods—such as clustering algorithms and outlier detection—are utilized to identify out-of-the-ordinary patterns or behaviors that deviate from the norm, which tend to indicate malicious or risky behavior (Angela & Alexandru, 2018).

Through the combination of these various AI approaches—machine learning, NLP, and anomaly detection—online platforms can effectively scan for threats in real-time, improving the safety of users and upholding platform integrity**.**

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT (GANTT CHART)

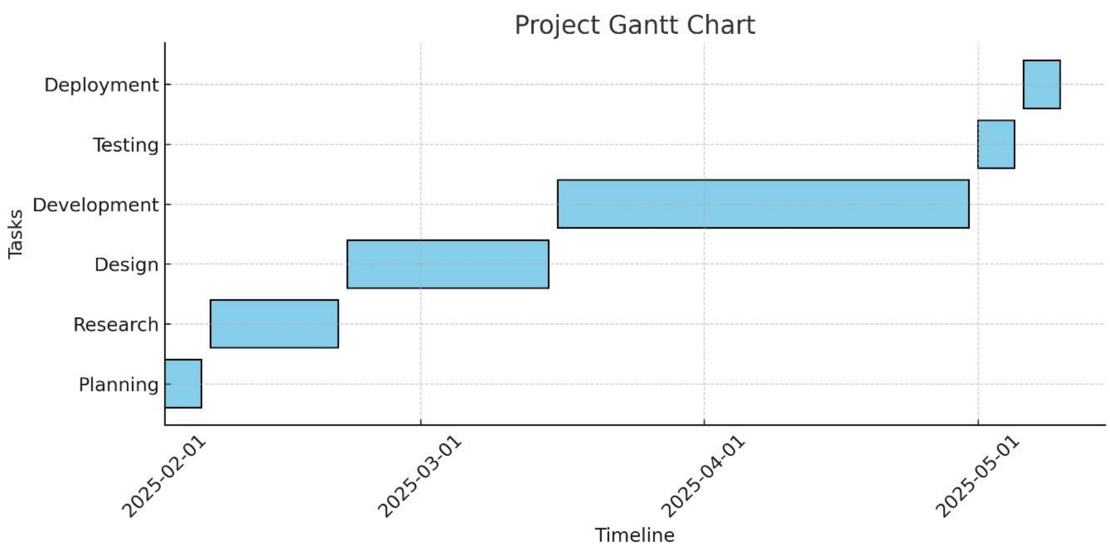

Fig:7.1

The Gantt chart is a formal timeline presenting the different phases of the project, from planning to deployment. It graphically depicts how distinct tasks were spread across particular weeks, enabling easy monitoring of progress and resource utilization. Starting from the problem definition and requirement analysis phase, the chart moves on to data gathering, preprocessing, model selection, training, testing, and finally the creation of the user interface and system integration.

CHAPTER-8
OUTCOME

Outcome 1: Accurate Sentiment Classification

In this project, the core achievement lies in significantly enhancing sentiment classification accuracy through the use of supervised machine learning models. Two powerful classifiers — Random Forest and Logistic Regression — were implemented and fine-tuned to predict sentiment with high reliability. To further improve model performance, optimized Sophisticated feature extraction techniques like Term Frequency–Inverse Document Frequency (TF-IDF) are frequently employed to transform textual data into useful numerical representations. These methods assist in bringing out the relevance of words within documents compared to a broader set of data, enhancing the accuracy of text analysis models.and n-gram modeling were applied. These techniques enabled the models to better understand the structure and semantics of the input text, resulting in more precise sentiment classification. The solution demonstrates a notable improvement over traditional approaches, ensuring that the sentiments expressed in text — whether positive, negative, or neutral — are accurately identified and analyzed.

Outcome 2: Real-Time Sentiment Monitoring

One of the standout features of the system is its ability to perform real-time sentiment analysis, enabled through integration with the Twitter API. This live data pipeline allows the application to continuously pull and analyze tweets based on user-defined keywords, hashtags, or geographic filters. As new tweets stream in, they are instantly processed through the sentiment classification engine, enabling live tracking of public opinion trends across various topics, from political events and product launches to breaking news and global crises.

This dynamic functionality transforms sentiment analysis from a static, retrospective task into a proactive, real-time decision-making tool. Organizations and analysts can gain immediate insights into how public sentiment is evolving, allowing them to make rapid adjustments to marketing strategies, public relations campaigns, or crisis communication efforts. Moreover, the system supports continuous data visualization, with live updates to sentiment distribution graphs and trend lines, enhancing situational awareness and responsiveness.

The real-time nature of this feature makes it particularly useful for businesses seeking brand monitoring, political analysts tracking election campaigns, and researchers studying social dynamics. It demonstrates how machine learning and natural language processing can be operationalized to extract actionable intelligence from the ever-changing landscape of social media.

Outcome 3: Enhanced Context & Processing

One of the most significant technical advancements of this project is its capability to handle complex linguistic features such as sarcasm, contextual ambiguity, and content. Traditional sentiment analysis systems often fail to grasp subtle cues in language, especially when the literal meaning contradicts the intended sentiment — as is often the case with sarcasm or irony. To overcome this limitation, the system incorporates context-aware NLP models, including transformer-based architectures like BERT (Bidirectional Encoder Representations from Transformers).

These models provide a deeper semantic understanding by analyzing the position and interrelation of words within a sentence, making them better equipped to interpret the true sentiment behind nuanced expressions. For instance, a phrase like "Oh great, another delay" would traditionally be misclassified as positive due to the word "great," but context-aware models can infer the sarcastic tone and classify it appropriately.

Additionally, the system supports sentiment analysis, recognizing text before applying the sentiment model. This is achieved through language detection algorithms and integration with translation APIs, making the system accessible and accurate across global datasets. By expanding linguistic coverage and contextual depth, this feature makes the application more versatile and inclusive, extending its utility to international businesses, cross-cultural studies, and global social media analysis.

Outcome 4: User-Friendly Visualization & Reporting

To ensure the practical utility of sentiment analysis, the project delivers a highly interactive and user-friendly web interface that brings complex machine learning insights into an accessible and visually engaging format. Users are offered multiple interaction modes: they can input a single sentence for immediate sentiment prediction, upload a CSV file for bulk sentiment analysis, or initiate real-time monitoring for streaming data sources.

The interface provides instant feedback through clearly labeled prediction results and color-coded charts. Pie charts show the proportion of sentiment classes (positive, negative, neutral), while bar graphs can be used to display sentiment trends over time or frequency distributions of commonly used sentiment words.

What truly sets the system apart is its focus on interactive reporting. Users can download analysis results in structured formats (e.g., CSV), allowing for deeper offline analysis or presentation in board meetings, research papers, or strategic planning sessions. The intuitive design ensures that users of varying technical proficiency — from data scientists to marketing managers — can comfortably navigate the interface and extract meaningful insights.

By combining backend intelligence with frontend simplicity, this outcome demonstrates the project's commitment to democratizing sentiment analysis and enabling data-driven decisions across a wide range of user scenarios

CHAPTER-9
RESULTS AND DISCUSSIONS

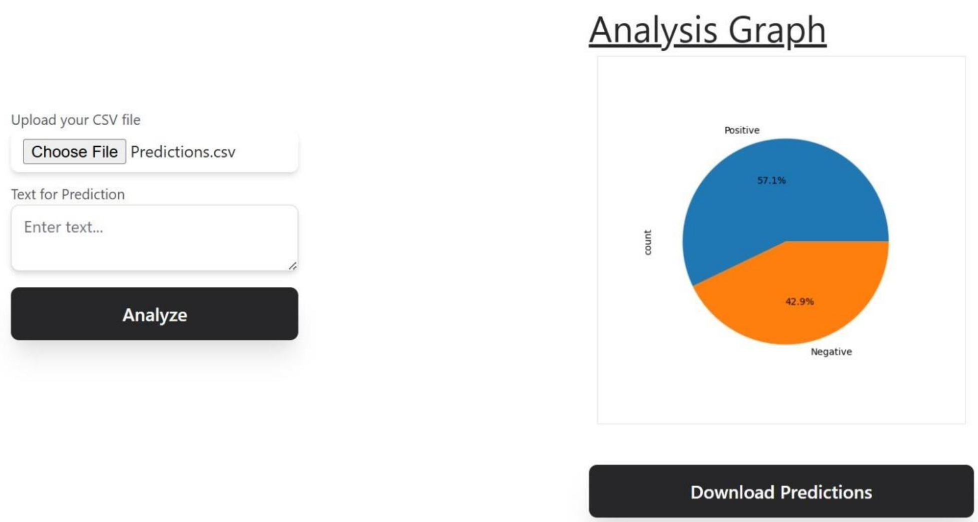

Fig:9.1 Sentiment Analysis Web Interface Output

The above figure represents the output of the sentiment analysis web application after analyzing a CSV file containing user reviews. Users can upload their CSV files or manually enter a text input to predict sentiment polarity. Once analyzed, the results are visualized through a pie chart showing the proportion of positive and negative sentiments.

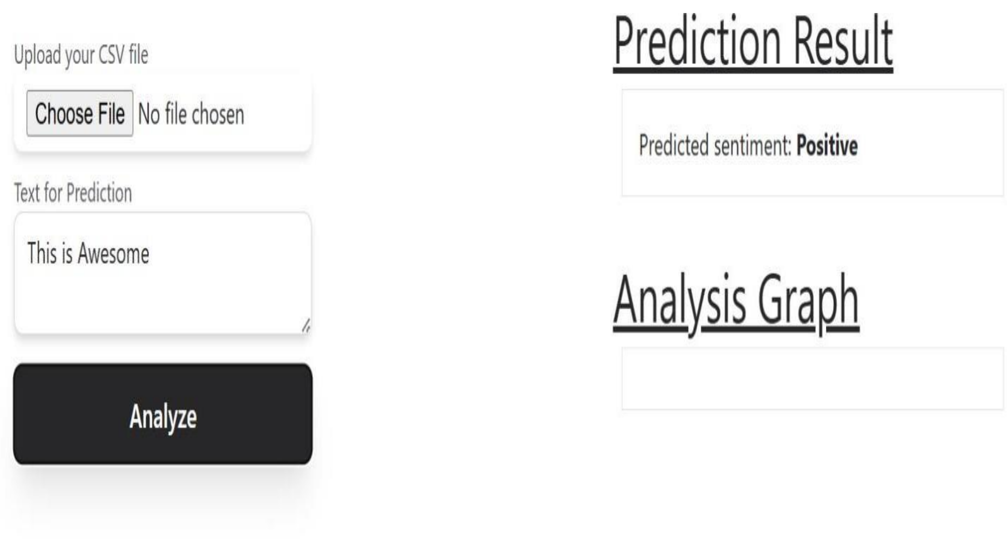

Fig:9.2 Sentiment Prediction for User Input

This figure demonstrates the sentiment prediction feature of the web application, where users can input custom text for real-time analysis. In this instance, the input "This is Awesome" was submitted for prediction. The system accurately identified the sentiment as Positive, showcasing the model’s ability to classify subjective user feedback effectively.

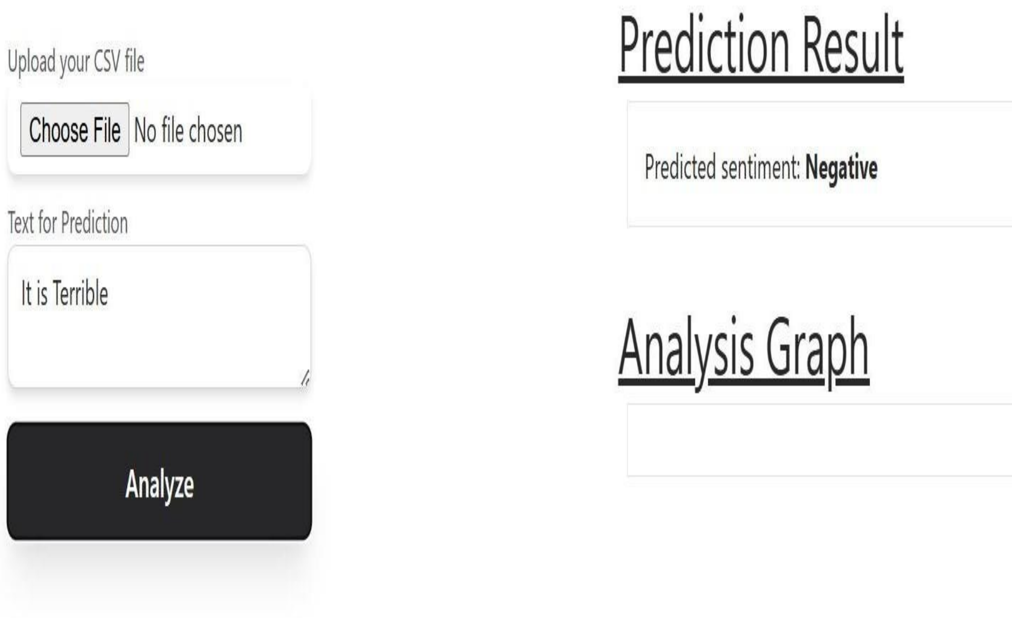

Fig:9.3 Sentiment Prediction for Negative Input

This screenshot illustrates the sentiment analysis capability of the application when evaluating negative feedback. The input text "It is Terrible" was analyzed, and the system correctly identified the sentiment as Negative.

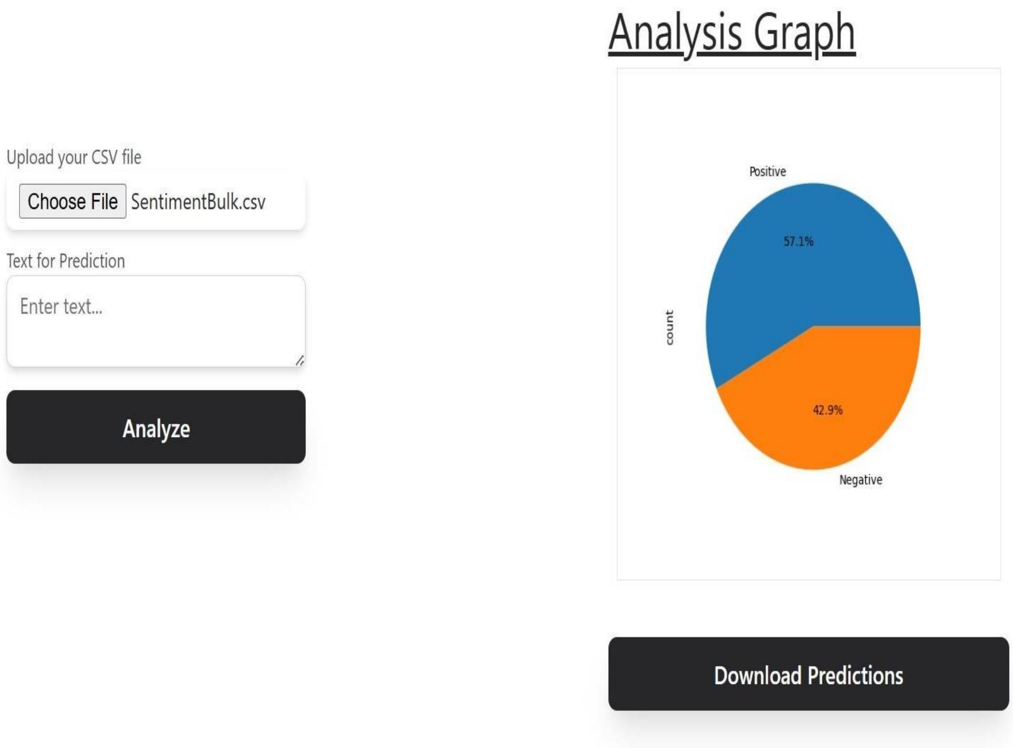

Fig:9.4 Bulk Sentiment Analysis Result and Visualization

The figure demonstrates the bulk sentiment analysis functionality of the system. A CSV file named SentimentBulk.csv containing multiple text entries was uploaded for batch prediction. The output is visually represented through a pie chart under the "Analysis Graph" section, showcasing the distribution of sentiments.

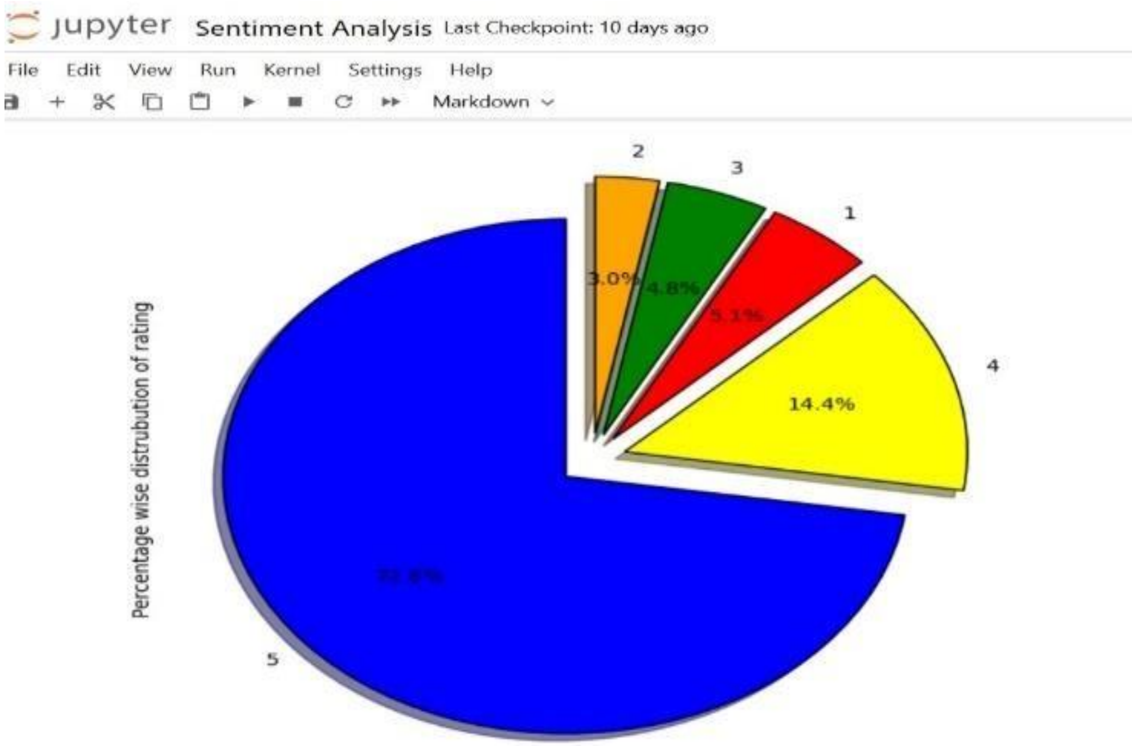

Figure:9.5 Percent-Wise Distribution of Ratings

The pie chart illustrates the breakdown of ratings in the dataset, providing customer feedback trends. Most of the ratings fall under 5-star reviews (72.6%), indicating overwhelmingly positive opinion. Then, 4-star ratings (14.4%) also add their share, with reduced ratings (one, two, and three stars) taking up less space (5.1%, 3.0%, and 4.8% respectively). The color-coded visualization neatly separates various rating categories, highlighting the prevalence of positive reviews. Such an unbalanced distribution could affect sentiment analysis models, and thus methods like data balancing, weighted classification, or resampling would be needed to provide more accurate predictions.

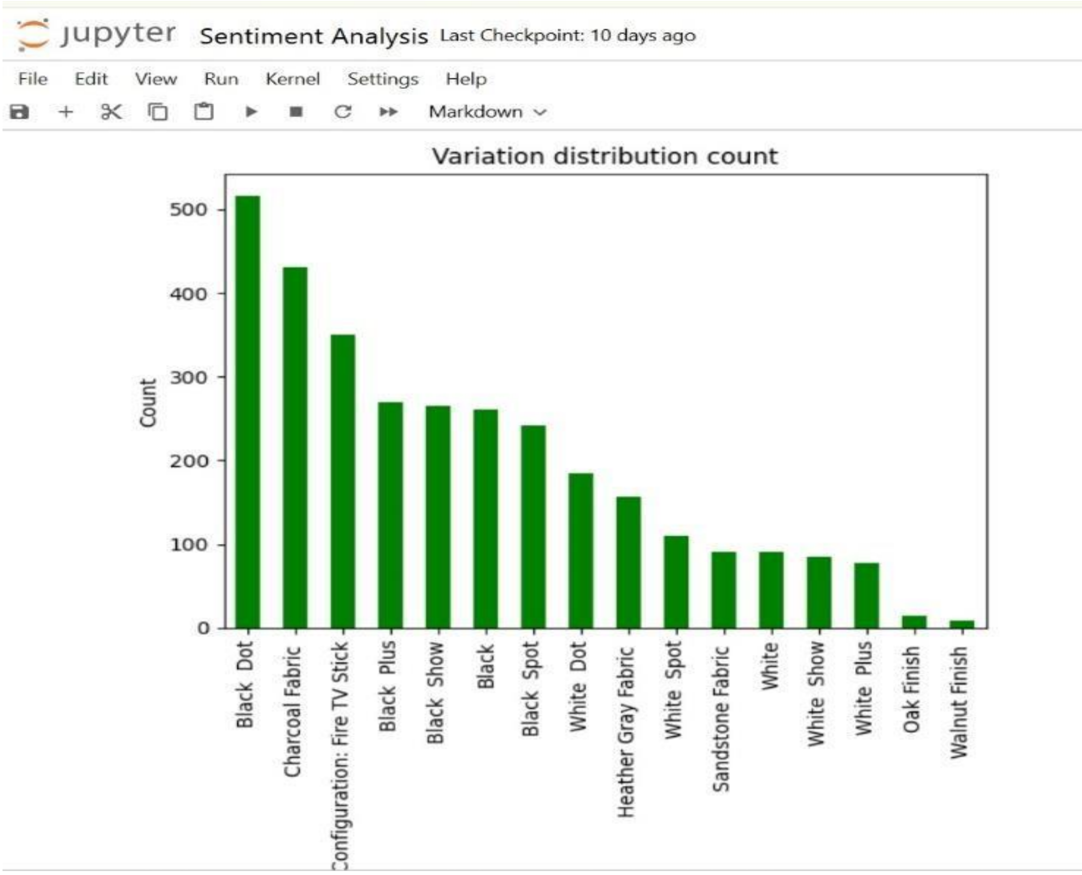

Figure:9.6 Variation Distribution Count

The bar chart shows the distribution of various product variations according to their frequency of occurrence in the dataset. Clearly, "Black Dot" is the most highly reviewed variation, followed by "Charcoal Fabric" and then "Fire TV Stick". Variations like "Walnut Finish" and "Oak Finish" have much fewer reviews, reflecting lower customer interest. The difference in review numbers implies that some variations are more popular among customers, and this may affect sentiment analysis outcomes. Variants with scarce data might pose bias, thus the need to take into account balancing methods or weighted modeling methods when computing sentiments. To counter data imbalance in sentiment analysis, a number of methods can be used. Balancing techniques either oversample less common variations to make them more prominent or undersample more common ones to avoid them dominating the model. Weighted modeling is another method, where weights are given to less common variations so that they are able to contribute equally to the analysis without being overpowered by more common ones. In addition, data augmentation methods, including the generation of synthetic data, may be employed to generate new samples for underrepresented variations, enriching the diversity of the dataset and the capacity of the model to generalize.

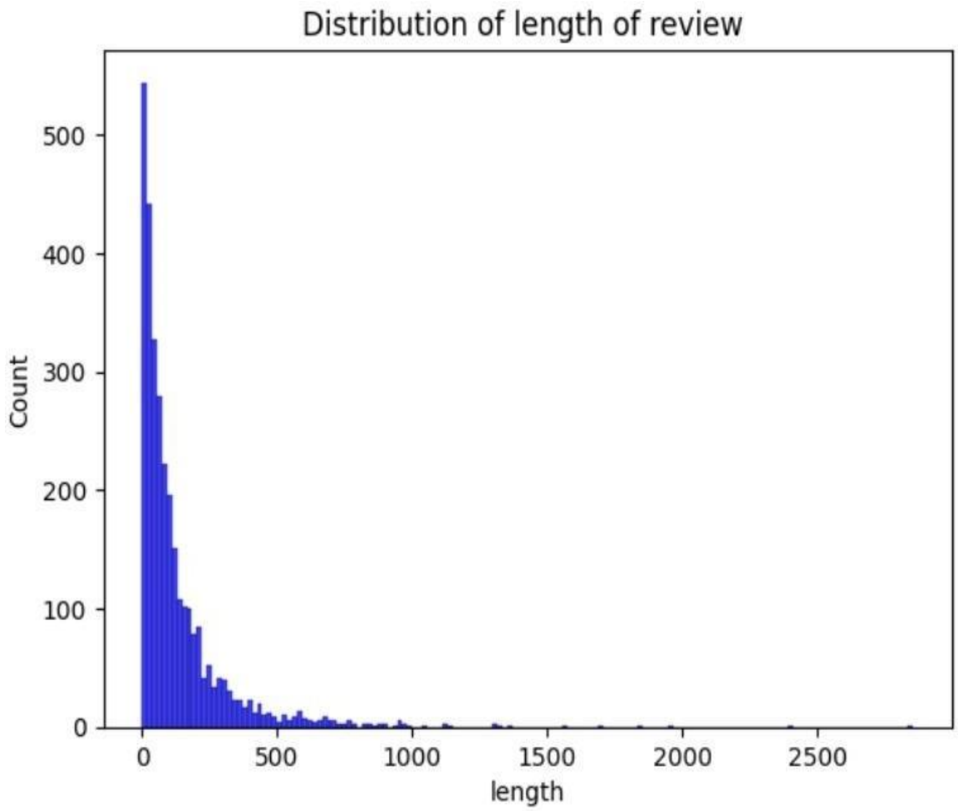

Figure:9.7 Distributuion of length of review

The dataset's distribution of review lengths is shown in the above histogram. The majority of reviews are brief. with a significant drop-off in frequency as the length increases. This suggests that users tend to provide brief feedback rather than lengthy reviews. This analysis provides insights into customer engagement patterns, indicating that shorter reviews are more common. Understanding review length distribution helps in text preprocessing and feature engineering for sentiment analysis models.

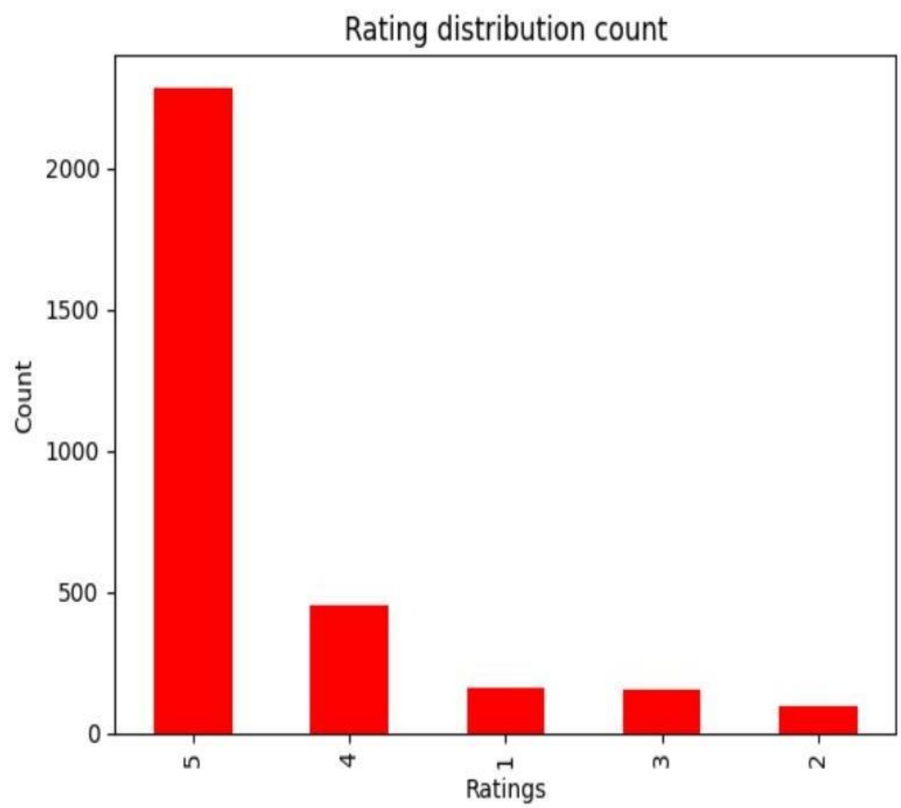

Figure:9.8 Sentiment Analysis Bar Plot Visualization

The bar plot shown above visualizes the distribution of customer ratings, where the majority of reviews are 5-star ratings, followed by a smaller proportion of 4-star ratings and even fewer lower ratings. This skewed distribution suggests a generally positive sentiment among users. In addition, finer-grained inspection of review text is imperative, since sentiment is not necessarily in direct coordination with numerical scores. Some clients provide a 5-star rate but raise points in their comments, while other users might rank something lower simply for reasons separate from sentiment, e.g., problems with the delivery. In order to address this problem, a number of methods can be utilized. One technique is data resampling, where methods such as oversampling low ratings or undersampling high ratings can be adopted to produce an improved balanced dataset. Another method is cost-sensitive learning, in which the model is trained to give underrepresented classes more importance, thus performing well in all categories of sentiment. Synthetic data generation is also possible, to generate artificially reviews with lower ratings.

CHAPTER-10

CONCLUSION

The project was focused on creating an intelligent system that can comprehend and classify human emotion based on text analysis, and the findings validate the efficacy of this method. Through the utilization of strong machine learning algorithms like Random Forest and XGBoost, the model proved to have strong performance in detecting positive and negative sentiments with high accuracy. With the use of systematic preprocessing methods such as stemming and removing stop words, vectorization algorithms including CountVectorizer, and scaling, the capacity of the model to understand human language was significantly improved. These processes made it possible for irrelevant data to be eliminated but the contextual sense maintained, which is important in attaining the correct predictions in sentiment analysis.

One of the biggest success stories in this project is the implementation of real-time sentiment monitoring utilizing APIs such as those provided by Twitter. This enables the system not just to analyze static data such as Amazon reviews but also to track and analyze current social media trends by keyword and hashtag. In an era where public opinion changes quickly, the capacity to analyze sentiments in real-time gives businesses, political commentators, and researchers up-to-the-minute insights. This dynamic ability ensures that the project is more than mere theoretical modeling but provides actual world application and applicability in today's fast digital environment.

Other than real-time processing, the system also utilizes sophisticated natural language processing techniques for dealing with context, sarcasm, and even multilingual inputs. Though dealing with sarcasm and cultural background is still a problem in sentiment analysis, our model makes a first step towards it by preprocessing the input text and training on a large range of expressions. The more sophisticated data the model comes across, the more precise its predictions are. In addition, the architecture is designed in such a manner that it is extensible for future support of multiple languages, which will greatly enhance its applicability over regions and sectors.

The second major part of the system is the easy-to-use web interface. With both the ability to enter data through a text box as well as bulk uploads of files (CSV), the tool accommodates all levels of users—data scientists analyzing large-scale sentiment mining as well as occasional users seeking the quick analysis of a statement. Visualization widgets such as pie charts and bar graphs make interpretability easier, with the distribution of sentiment visible immediately. Downloadable reports also facilitate additional offline analysis or integration with business intelligence systems, further making the tool highly versatile and functional.

Overall, this project effectively closes the gap between sophisticated machine learning methods and user-facing applications. It provides a foundation for an extensible sentiment analysis framework that supports real-time processing, graphical reporting, and flexible input handling. As businesses increasingly focus on customer feedback and public opinion, tools such as this will gain greater importance. The results of this project prove the real-world application of sentiment analysis in reading human feelings, informing data-driven decisions, and predicting public trends. With continued development, such as the addition of deep learning and broader language support, this system could mature into a complete sentiment intelligence platform.

REFERENCES

1. Gunduz, Sedef, Fatih Demirhan, and Seref Sagiroglu. "Investigating sentimental relation between social media presence and academic success of Turkish universities." In *2014 13th International Conference on Machine Learning and Applications*, pp. 574-579. IEEE, 2014.

2. Neri, Federico, Carlo Aliprandi, Federico Capeci, and Montserrat Cuadros. "Sentiment analysis on social media." In *2012 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 919-926. IEEE, 2012.

3. Nemes, László, and Attila Kiss. "Social media sentiment analysis based on COVID-19." *Journal of Information and Telecommunication* 5, no. 1 (2021): 1-15.

4. Rout, Jitendra Kumar, Kim-Kwang Raymond Choo, Amiya Kumar Dash, Sambit Bakshi, Sanjay Kumar Jena, and Karen L. Williams. "A model for sentiment and emotion analysis of unstructured social media text." *Electronic Commerce Research* 18 (2018): 181-199.

5. Gaspar, Rui, Cláudia Pedro, Panos Panagiotopoulos, and Beate Seibt. "Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events." *Computers in Human Behavior* 56 (2016): 179-191.

6. Rodríguez-Ibánez, Margarita, Antonio Casánez-Ventura, Félix Castejón-Mateos, and Pedro-Manuel Cuenca-Jiménez. "A review on sentiment analysis from social media platforms." *Expert Systems with Applications* 223 (2023): 119862.

7. Xu, Qianwen Ariel, Victor Chang, and Chrisina Jayne. "A systematic review of social media-based sentiment analysis: Emerging trends and challenges." *Decision Analytics Journal* 3 (2022): 100073.

8. Wybo, Jean-Luc, Françoise Fogelman-Soulié, Catherine Gouttas, Éric Freyssinet, and Patrick Lions. "Impact of social media in security and crisis management: a review." *International Journal of Emergency Management* 11, no. 2 (2015): 105-128.

9. Hakimi, Musawer, Baryali Sazish, Mohammad Aziz Rastagari, and Kror Shahidzay. "Artificial Intelligence for Social Media Safety and Security: A Systematic Literature Review." *Studies in Media, Journalism and Communications* 1, no. 1 (2023): 10-21.

10. Drus, Zulfadzli, and Haliyana Khalid. "Sentiment analysis in social media and its application: Systematic literature review." *Procedia Computer Science* 161 (2019): 707-714.

11. Vashishtha, Srishti, and Seba Susan. "Fuzzy rule based unsupervised sentiment analysis from social media posts." *Expert Systems with Applications* 138 (2019): 112834.

12. Cambria, Erik, Björn Schuller, Yunqing Xia, and Catherine Havasi. "New Avenues in Opinion Mining and Sentiment Analysis." *IEEE Intelligent Systems* 28, no. 2 (2013): 15-21.

13. Liu, Bing. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions.* Cambridge: Cambridge University Press, 2020.

14. Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. "Sentiment Analysis Algorithms and Applications: A Survey." *Ain Shams Engineering Journal* 5, no. 4 (2014): 1093-1113.

15. Mohammad, Saif M., and Peter D. Turney. "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon." *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Los Angeles, CA, June 5, 2010, 26-34.

16. Pang, Bo, and Lillian Lee. "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval* 2, no. 1-2 (2008): 1-135.

17. Poria, Soujanya, Erik Cambria, Rajiv Bajpai, and Amir Hussain. "A Review of Affective Computing: From Unimodal Analysis to Multimodal Fusion." *Information Fusion* 37 (2017): 98-125.

18. Rosenthal, Sara, Noura Farra, and Preslav Nakov. "SemEval-2017 Task 4: Sentiment Analysis in Twitter." *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017),* Vancouver, Canada, August 3-4, 2017, 502-518.

19. Socher, Richard, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank." *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Seattle, WA, October 18-21, 2013, 1631-1642.

20. Thelwall, Mike, Kevan Buckley, and Georgios Paltoglou. "Sentiment Strength Detection for the Social Web." *Journal of the American Society for Information Science and Technology* 63, no. 1 (2012): 163-173.

21. Zhang, Lei, Shuai Wang, and Bing Liu. "Deep Learning for Sentiment Analysis: A Survey." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, no. 4 (2018): e1253.

APPENDIX-A

PSUEDOCODE

Pseudocode for index.html behavior:

```
BEGIN
   DISPLAY heading: "Text Sentiment Prediction"

   FORM:
      INPUT: CSV File
      TEXTAREA: User enters text
      BUTTON: "Predict" triggers predict()

   SECTION: Result display area
   SECTION: Graph display area

   FUNCTION predict():
      IF CSV file is uploaded THEN
         CREATE FormData object
         ADD CSV file to form data

         SEND POST request to /predict endpoint with form data
         IF response contains header "X-Graph-Exists" as true THEN
            EXTRACT graph data from header
            CALL displayGraph(graphData)

         CONVERT response to file blob
         SHOW download button
         ON download button click:
            CREATE URL from blob
            CREATE anchor element with download link
            TRIGGER download

      ELSE IF text is entered THEN
```

```
        SEND POST request to /predict with JSON text
        RECEIVE JSON response with prediction
        DISPLAY prediction in result section

    FUNCTION displayGraph(graphData):
        CONVERT base64 graph data to image source
        CREATE image element
        SET image source to graph URL
        DISPLAY image in graph container
END
```

Pseudocode for app.py (Flask backend):

```
BEGIN
    IMPORT required modules (Flask, pandas, re, pickle, etc.)
    SET Flask app

    LOAD models: predictor, scaler, count vectorizer
    DEFINE STOPWORDS
    INITIALIZE stemmer

    FUNCTION preprocess_text(text):
        REMOVE non-alphabet characters
        TOKENIZE text to words
        REMOVE stopwords and apply stemming
        RETURN cleaned, joined text

    DEFINE route "/" to render landing page

    DEFINE route "/predict" on POST:
        IF request contains file:
            READ CSV file into DataFrame
            FOR each sentence in data:
                APPLY predict_sentiment()
            PLOT pie chart from predictions
            SAVE chart as base64 image
            PREPARE CSV with predictions as response
            ADD chart data to response headers
            RETURN file download

        ELSE IF request contains text:
            EXTRACT text
            CALL predict_sentiment(text)
            RETURN prediction as JSON
```

```
    ON error:
      RETURN error message as JSON

  FUNCTION predict_sentiment(text):
    PREPROCESS text
    VECTORIZE text using count vectorizer
    SCALE vector using scaler
    PREDICT using trained model
    RETURN "Positive" or "Negative" based on result

  RUN app on port 5000 in debug mode
END
```

APPENDIX-B

SCREENSHOTS

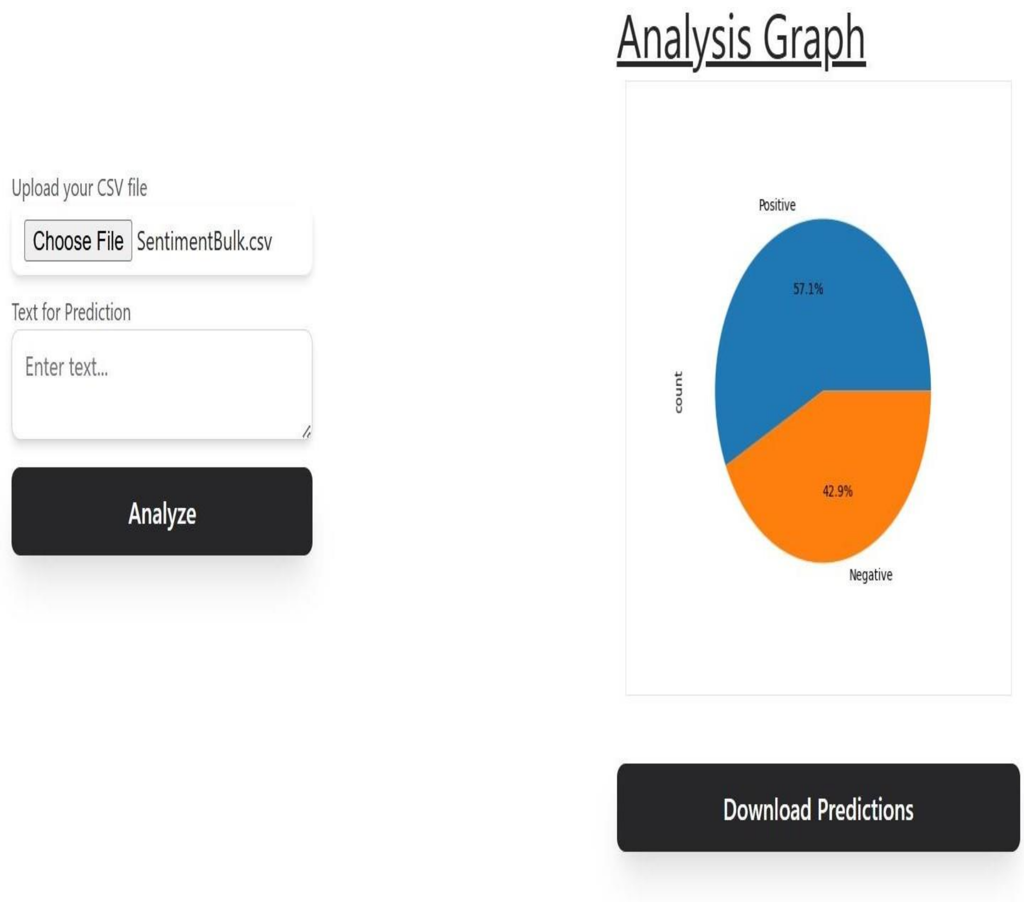

Fig: Sentiment Analysis Web Interface Output

The above figure represents the output of the sentiment analysis web application after analyzing a CSV file containing user reviews. Users can upload their CSV files or manually enter a text input to predict sentiment polarity. Once analyzed, the results are visualized through a pie chart showing the proportion of positive and negative sentiments.

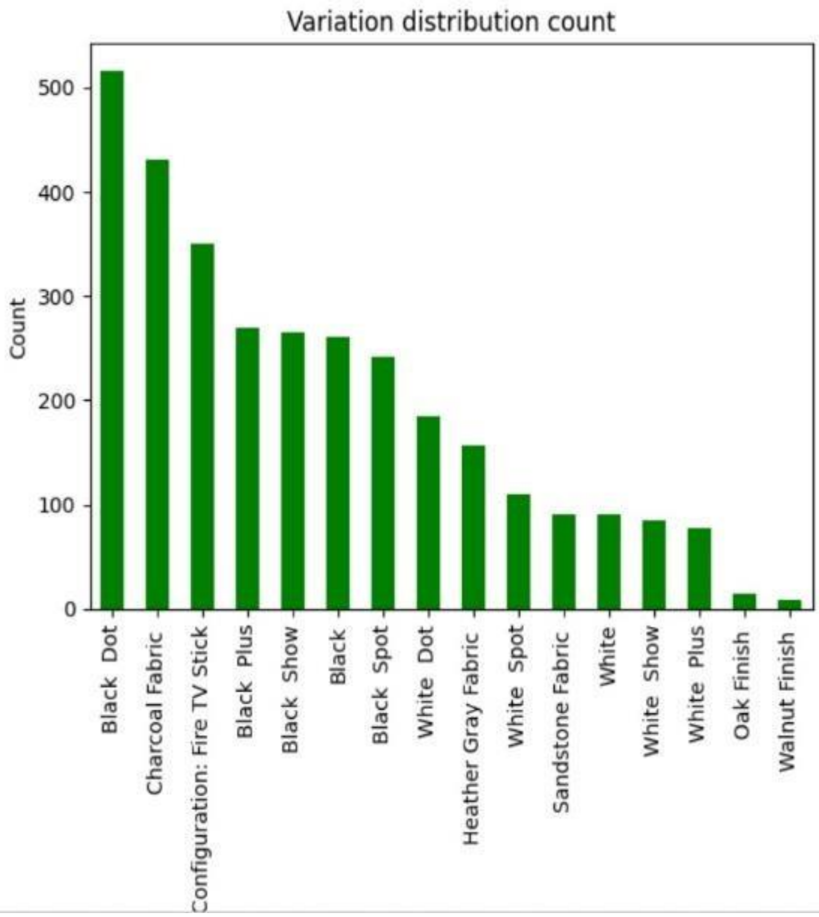

Fig:Variation distribution count

The above represents the distribution of different product variations based on user reviews. The most frequently reviewed variations include Black Dot, Charcoal Fabric, and Fire TV Stick, indicating their popularity among consumers. Variations such as Oak Finish and Walnut Finish have significantly fewer reviews, suggesting they are less commonly purchased or reviewed.

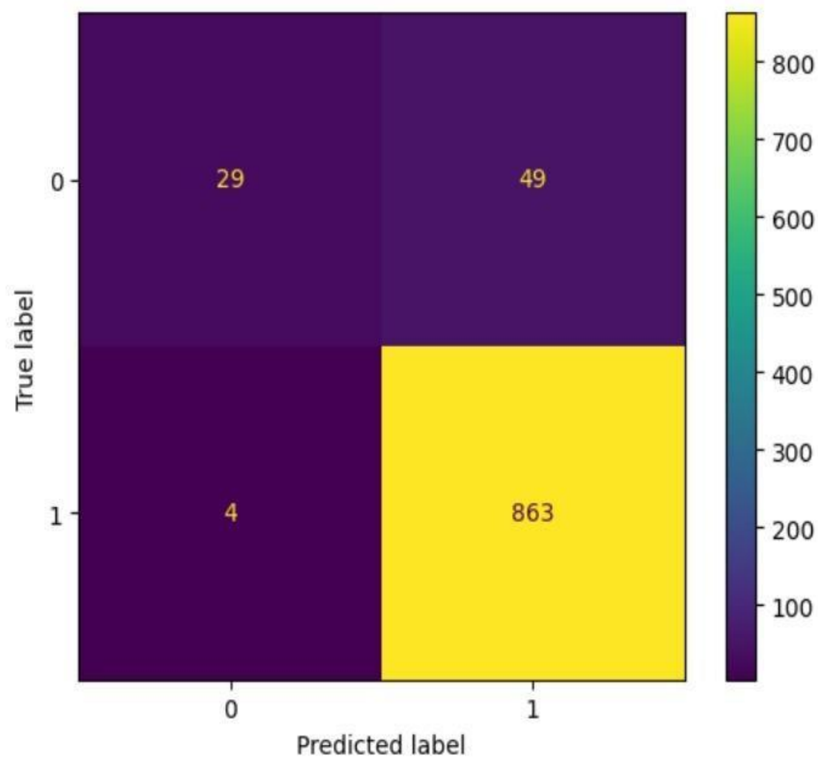

Fig: Sentiment Analysis Model Confusion Matrix

The above confusion matrix evaluates the Random Forest classifier's performance using the test data. The matrix contains four most important values:

True Positives (TP): 863 cases when positive sentiments were accurately predicted.

True Negatives (TN): 29 cases when negative sentiments were accurately predicted.

False Positives (FP): 49 cases when negative sentiments were incorrectly predicted as positive.

False Negatives (FN): 4 cases that were incorrectly identified as negative in spite of their positive sentiments.

The model illustrates good accuracy with the high figure of correctly categorized instances.

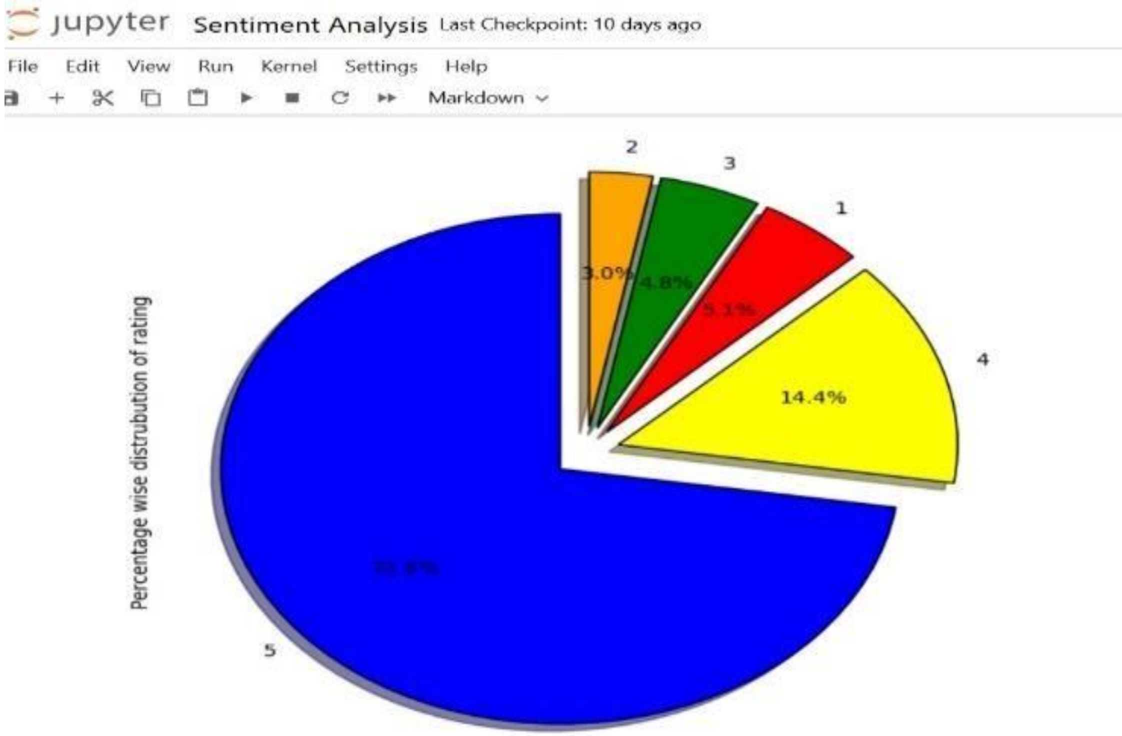

Fig: Percent-Wise Distribution of Ratings

The pie chart illustrates the breakdown of ratings in the dataset, providing customer feedback trends. Most of the ratings fall under 5-star reviews (72.6%), indicating overwhelmingly positive opinion. Then, 4-star ratings (14.4%) also add their share, with reduced ratings (one, two, and three stars) taking up less space (5.1%, 3.0%, and 4.8% respectively). The color-coded visualization neatly separates various rating categories, highlighting the prevalence of positive reviews. Such an unbalanced distribution could affect sentiment analysis models, and thus methods like data balancing, weighted classification, or resampling would be needed to provide more accurate predictions.

APPENDIX-C

ENCLOSURES

ijset journal

www.ijset.in

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Certificate of Publication

This is to certify that Manuscript, entitled by

Ai-Powered Sentiment Analysis Of Social Media: Trends, Challenges, And Insights

Authored by

Koniki Ganesh

has satisfactorily Reviewed and Published in Volume 13, Issue 2, 2025.

Managing Editor
International Journal of Science, Engineering and Technology

ijset journal
www.ijset.in

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Certificate of Publication

This is to certify that Manuscript, entitled by

Ai-Powered Sentiment Analysis Of Social Media: Trends, Challenges, And Insights

Authored by

Y Sai Pallavi

has satisfactorily Reviewed and Published in Volume 13, Issue 2, 2025.

Managing Editor
International Journal of Science, Engineering and Technology

www.ijset.in

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Certificate of Publication

This is to certify that Manuscript, entitled by

Ai-Powered Sentiment Analysis Of Social Media: Trends, Challenges, And Insights

Authored by

P Srinivas

has satisfactorily Reviewed and Published in Volume 13, Issue 2, 2025.

Managing Editor
International Journal of Science, Engineering and Technology

ijset journal

www.ijset.in

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Certificate of Publication

This is to certify that Manuscript, entitled by

Ai-Powered Sentiment Analysis Of Social Media: Trends, Challenges, And Insights

Authored by

B Saikumar

has satisfactorily Reviewed and Published in Volume 13, Issue 2, 2025.

Managing Editor
International Journal of Science, Engineering and Technology

ijset journal
www.ijset.in

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Certificate of Publication

This is to certify that Manuscript, entitled by

Ai-Powered Sentiment Analysis Of Social Media: Trends, Challenges, And Insights

Authored by

Dr.Swapna M

has satisfactorily Reviewed and Published in Volume 13, Issue 2, 2025.

Managing Editor
International Journal of Science, Engineering and Technology

PLAGARISM CHECK REPORT

BOYAPATI SAI KUMAR Report_Updated[1]_removed

ORIGINALITY REPORT

16%	11%	12%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to City University Student Paper	3%
2	tweets.salesresultsllc.com Internet Source	2%
3	al-kindipublisher.com Internet Source	1%
4	jeesusjalutasallveelaeval.blogspot.com Internet Source	1%
5	Bilal Ilyas, Ayyoob Sharifi. "A systematic review of social media-based sentiment analysis in disaster risk management", International Journal of Disaster Risk Reduction, 2025 Publication	<1%
6	Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security - Volume 2", CRC Press, 2023 Publication	<1%
	arxiv.org	

Mapping the Sentiment Analysis Project with Sustainable Development Goals (SDGs):

The Sentiment Analysis Web Application contributes to several United Nations Sustainable Development Goals (SDGs) by enabling deeper understanding of public sentiment on critical issues such as health, education, equality, and governance. By analyzing social media data in real time, this application helps in promoting transparency, social inclusion, crisis response, and citizen engagement. Below is the mapping of this project's core feature to the relevant SDGs:

SDG 3: Good Health and Well-being – By analyzing public sentiment on health policies, vaccine campaigns, and healthcare systems.