CMR ENGINEERING COLLEGE
UGC AUTONOMOUS
Approved by AICTE | Affiliated to JNTUH | Accredited by NAAC & NBA

Department of Computer Science & Engineering(AI &ML)

A Mini Project Review-II Presentation
on

UTILIZING MACHINE LEARNING FOR AN EFFICIENT ANALYSIS OF CUSTOMERS PERSPECTIVE FROM TWEETS

Internal Guide
Mr.B.Balakrishna
Assistant Professor
Dept of CSE-AIML.

By
D.SAIRAM :218R1A66E6
A.RAKESH :218R1A66D4
K.PRAVEENKUMAR :218R1A66G0
HM.AKSHAYKUMAR :218R1A66F6

CONTENT

DOMAIN SELECTION

Why Machine Learning ?

- **Abundance of Data:** Social media generates tons of data. Machine learning helps us sift through this massive amount efficiently.

- **Global insights:** Twitter is global. Analyzing sentiments here provides a broad, international perspective, crucial in our interconnected world.

- **Rapid Feedback:** Twitter moves fast. Our domain allows us to capture and analyze sentiments in real-time, giving organizations an edge.

- **Wide Application:** By using machine learning, our project equips organizations to make informed decisions promptly, aligning with the fast-paced nature of social media dynamics.

WHY THIS TITLE?

- **ML as Tech:** Our title says we're using Machine Learning tech to handle lots of customer feedback on Twitter.
- **Customer's View Matters:** Project is all about understanding what customers think, so we can make products better based on what they say.
- **Decisions Made Easy:** Our project helps businesses make smart choices fast, using tech to keep up with what customers want.

Justification:

The title encapsulates the core objectives of our project - leveraging machine learning for a precise and efficient analysis of customer perspectives from tweets, aligning with the contemporary need for data-driven decision-making in the realm of social media.

ABSTRACT

- This project introduces a Twitter Sentiment Analysis system.
- The system utilizes a Python script employing the Tweepy and TextBlob libraries to extract, process, and analyze tweets from a pre-existing dataset.
- The sentiment analysis categorizes tweets as positive, negative, or neutral, and the results are presented through both textual output and dynamic visualizations.
- The system integrates Matplotlib for graphical representations, showcasing the distribution of sentiments through pie and bar charts.
- Additionally, the project explores the percentage distribution of positive, negative, and neutral tweets.
- Through a modular and comprehensible design, the proposed system provides insights into the sentiment landscape of Twitter data, offering a valuable tool for sentiment analysis enthusiasts and researchers.

INTRODUCTION

Social media is vital for opinions and daily experiences. Online communities like Twitter shape consumer perspectives and connect businesses directly. Organizations struggle with data extraction in such web-based applications.

Problem Statement:

- Extracting sentiment from massive volumes of informal language, emotions, and short-form communication

Objective:

- Develop a program for sentiment analysis of customer reviews on products from Twitter data.

SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS

Processor : Dual Core 2 duo

Hard disk : 250 GB

Ram : 2 GB DD Ram

SOFTWARE REQUIREMENTS

Front End : HTML, CSS, Python

Back End : MySql 5.5

Operating System : Windows 7

IDE : Spyder, Pycharm

SYSTEM REQUIREMENTS SPECIFIACTION

tweepy==4.6.0

textblob==0.17.1

numpy==1.19.5

matplotlib==3.3.1

EXISTING SYSTEM

Lexicon-Based Approach

Definition: In a lexicon-based approach, sentiment analysis relies on predefined dictionaries or lexicons that contain words associated with specific sentiment scores (positive, negative, or neutral).

Word Scoring: Each word in a text is assigned a sentiment score based on its presence in the lexicon. The overall sentiment of the text is determined by aggregating the scores of individual words.

Negation and Intensity: Some lexicon-based approaches consider negation (e.g., "not good") and intensity (e.g., "very good") to adjust sentiment scores accordingly.

Rule-Based: Lexicon-based approaches are often rule-based and involve simple algorithms for calculating sentiment scores.

PROPOSED SYSTEM

TextBlob ML Approach

Definition: TextBlob is a natural language processing (NLP) library in Python that provides a simple API for common NLP tasks, including sentiment analysis.

Algorithmic Approach: TextBlob uses a machine learning algorithm to analyze and classify the sentiment of a piece of text. It is based on a Naive Bayes classifier trained on a labeled dataset.

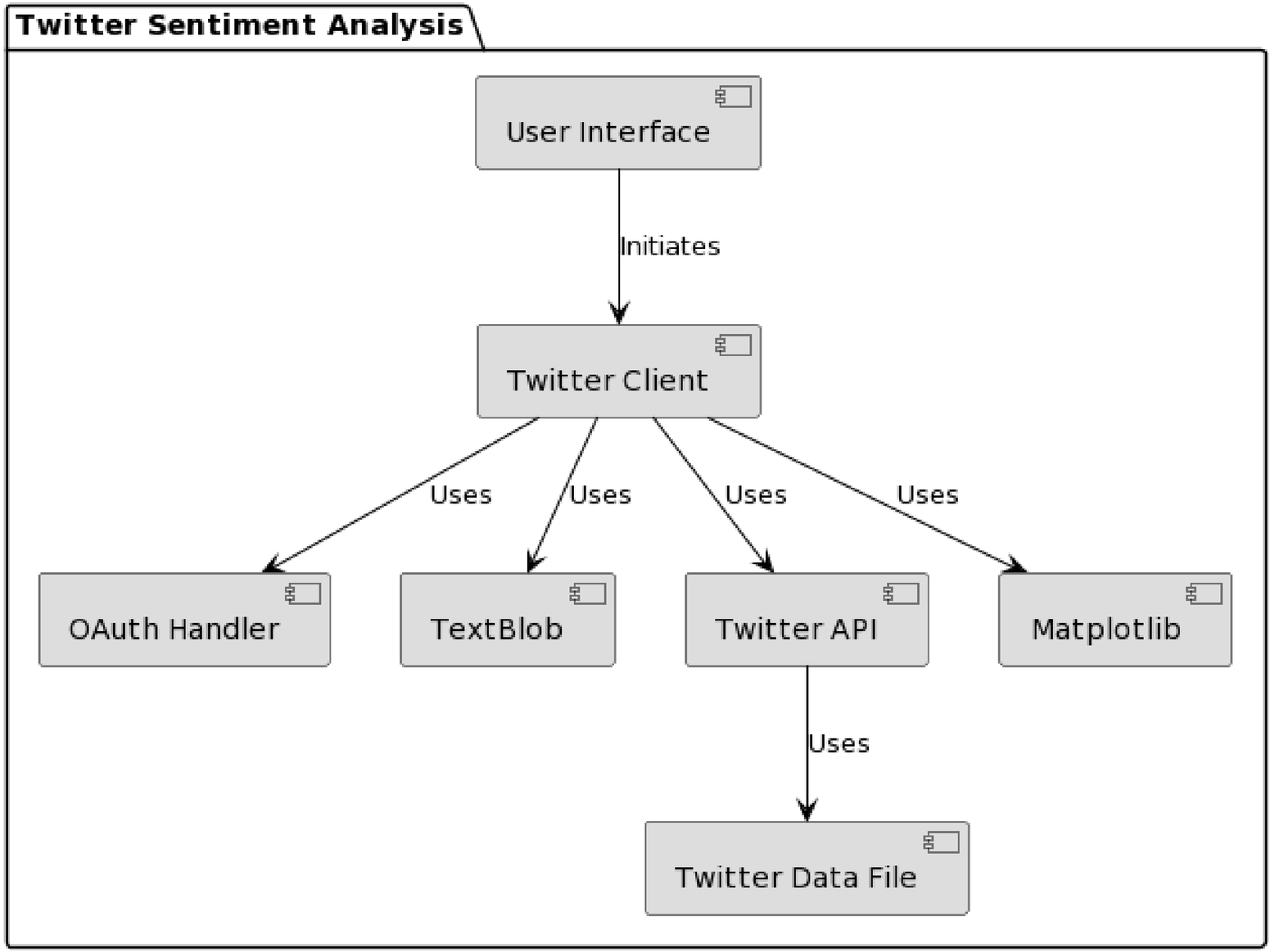
Pre-Trained Model: TextBlob comes with a pre-trained sentiment analysis model, which means it doesn't require explicit training on a specific dataset for general sentiment analysis tasks.

Strengths: It is suitable for a wide range of natural language processing tasks, and its sentiment analysis capabilities are more versatile than simple lexicon-based methods.

SYSTEM ANALYSIS & DESIGN

- Fetching and analyzing tweets, visualizing sentiment, and presenting results.
- Feasibility analysis confirms the project's operational, technical, and economic viability. System modeling, through use case, class, and sequence diagrams, provides a visual representation of the system's structure and interactions.
- Risk analysis identifies potential challenges, with mitigation strategies focused on updates and performance. Security measures ensure data integrity and proper authentication.
- Usability analysis emphasizes a clear interface and user training. Performance assessment focuses on execution times and scalability.
- Maintenance analysis underscores modularity and documentation for seamless updates.

In summary, the system analysis serves as a roadmap for the effective development and ongoing support of the Twitter Sentiment Analysis tool.



ARCHITECTURE

IMPLEMENTATION

1. TwitterClient Class:

- __init__: Initializes and authenticates Twitter API credentials.
- clean_tweet: Cleans tweet text for improved readability.
- get_tweet_sentiment: Classifies tweet sentiment using TextBlob.

2. Main Function:

- main: Orchestrates the program flow, fetching, cleaning, and analyzing tweets. Generates visualizations and prints results.

3. Visualization Functions:

- generate_pie_chart: Creates a pie chart for sentiment distribution.
- generate_bar_chart: Develops a bar chart for additional sentiment insights.

4. Performance and Scalability:

- optimize_execution_time: Optimizes sentiment analysis and visualization for faster execution.
- evaluate_scalability: Assesses system performance with varying dataset sizes.

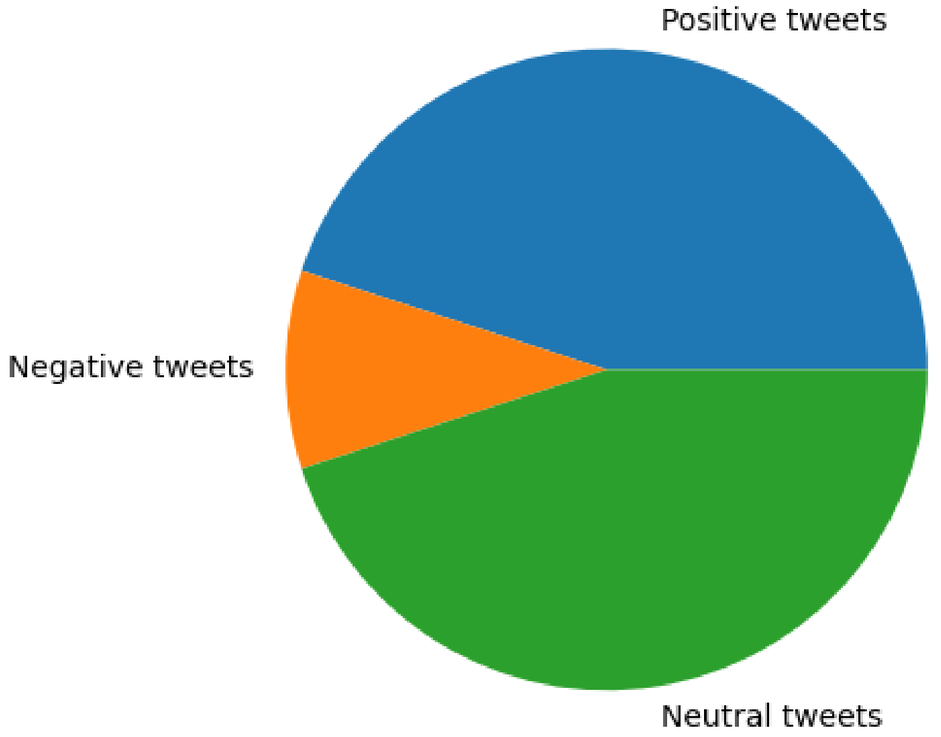
5. Documentation Functions:

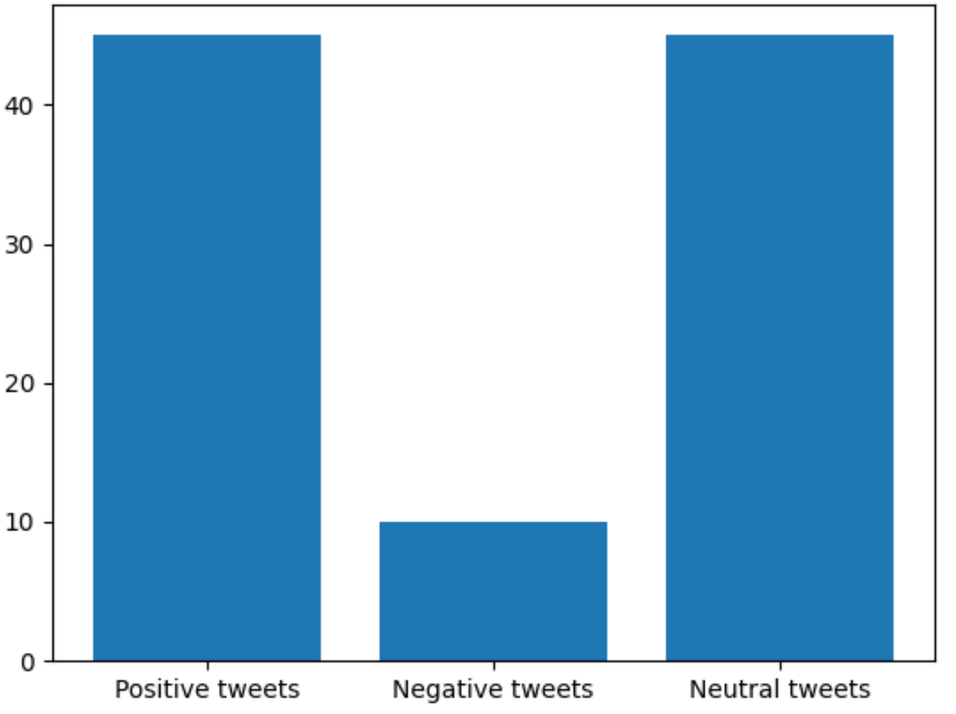
- generate_methods_documentation: Creates documentation for methods and algorithms.
- generate_results_documentation: Documents sentiment analysis outcomes.
- log_challenges: Logs and resolves development challenges.

OUTPUT:

```
PS C:\Users\raju espi\Desktop\CODE> python -u "c:\Users\raju espi\Desktop\CODE\Sentiment.py"
Positive tweets percentage: 45.0 %
Negative tweets percentage: 10.0 %
Neutral tweets percentage: 45.0 %
```

```
Positive tweets:
@chanx5268 @tiwarymanoj Former India Cricketer, opened with Kris Srikanth. Should have played more but those were d… https://t.co/xib0cfB689
RT @DHONIism: RIP VB Chandrasekar sir. 😔

The man who Selected Dhoni to INDIAN team in 2004 and Bought Dhoni to Chennai Super Kings in 2008…
RT @DHONIism: RIP VB Chandrasekar sir. 😔

The man who Selected Dhoni to INDIAN team in 2004 and Bought Dhoni to Chennai Super Kings in 2008…
RT @KKRidersFanClub: @Bazmccullum 's association with @KKRiders goes all the way back to the first ever match of the IPL in 2008 where he s…
RT @KKRidersFanClub: @Bazmccullum played for @KKRiders from 2008 to 2010 and also from 2012 to 2013. He was part of the title-winning side…
RT @KKRUniverse: McCullum returns to KKR, this time as a coach

Read more at: https://t.co/WLmuwXuj7T
#BAZisBack
RT @dailystarnews: Former New Zealand captain @Bazmccullum is appointed head coach of #Kolkata Knight Riders in the @IPL
https://t.co/6rsC…
RT @KKRUniverse: Former Knight Brendon McCullum is coming back to his first IPL team #KKR 😎🙌 But not as a player this time, Purple Army has…
RT @DHONIism: RIP VB Chandrasekar sir. 😔

The man who Selected Dhoni to INDIAN team in 2004 and Bought Dhoni to Chennai Super Kings in 2008…


Negative tweets:
This year:
GOT : Disappointed
IPL : Disappointed
WC : Disappointed
Now SG : Disappointed
Picking Jason Roy to open in an Ashes series was unfathomably stupid and wonderfully predictable #cricket #ipl2019… https://t.co/7zyAL3XXgn
PS C:\Users\raju espi\Desktop\CODE>
```


Positive tweets
Negative tweets
Neutral tweets


40
30
20
10
0
Positive tweets
Negative tweets
Neutral tweets

REFERENCES

01 M. Rambocas, and J. Gama, “Marketing Research:
The Role of Sentiment Analysis”. The 5th SNA-KDD Workshop’11. University of Porto, 2013.

02 A. K. Jose, N. Bhatia, and S. Krishna,
“TwitterSentimentAnalysis”. National Instituteof Technology Calicut, 2010.

03 P. Lai, “Extracting Strong Sentiment Trendfrom Twitter”. Stanford
University, 2012.

04 Y. Zhou, and Y. Fan, “ A Sociolinguistic Study of American Slang,” Theory and Practice in Language Studies, 3(12), 2209–2213, 2013. doi:10.4304/tpls.3.12.2209-2213

05 A.H.Huang, D.C. Yen, & X. Zhang, “Exploring the effects of emoticons,” Information & Management, 45(7), 466–473, 2008

Any Queries?

Thank You