A Mini Project Report on

"Twitter Sentiment Analysis on the Israeli–Palestinian conflict Using Vader Model"

Submitted in partial fulfillment of the requirement for the Sixth Semester

Bachelor of Engineering
In
Computer Science and Engineering
Visvesvaraya Technological University, Belgaum

<u>Submitted by</u>

Gauravdev HB , 1DS20CS073

Fahad HS, 1DS20CS067

Harsha S, 1DS20CS083

Ashrith reddy, 1DS20CS100

Under the guidance of
Prof Sarala
Designation,
Dept. of CSE, DSCE

2023-2024
Department of Computer Science and Engineering,
DAYANANDA SAGAR COLLEGE OFENGINEERING
BANGALORE – 560078

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM
DAYANANDA SAGAR COLLEGE OF ENGINEERING

Shavige Malleshwara Hills, Kumaraswamy Layout, Bangalore-560078
Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the mini project entitled "**Twitter Sentiment Analysis on the Israeli–Palestinian conflict using vader model**" is a bonafide work carried out by **"Gauravdev H B [1DS20CS073]"** in partial fulfilment of 7th semester, **Bachelor of Engineering** in **Computer Science and Engineering** under **Visvesvaraya Technological University, Belgaum** during the year 2023-24.

Prof Sarala   **Vice Principal & HOD**   **Principal,**
Prof/Ass.Prof/Assist Prof **Dept. of CSE, DSCE**   **DSCE,**
 CSE, DSCE, Bangalore. **Bangalore**   **Bangalore.**
Signature: _____ **Signature: _____** **Signature: _____**

Name of the Examiners   Signature with date
1………………………….   …………………………...
2………………………….   …………………………...

ACKNOWLEDGEMENT

Minor Project- Report
October 2023-December 2023

Course Faculty: Prof. Sarala
Semester:7th

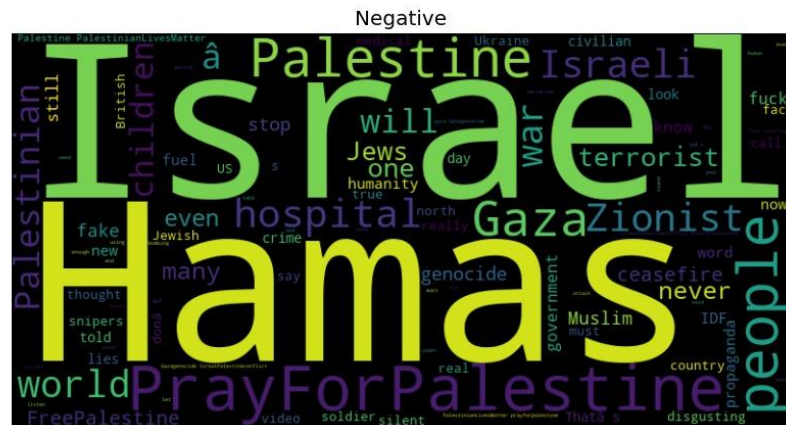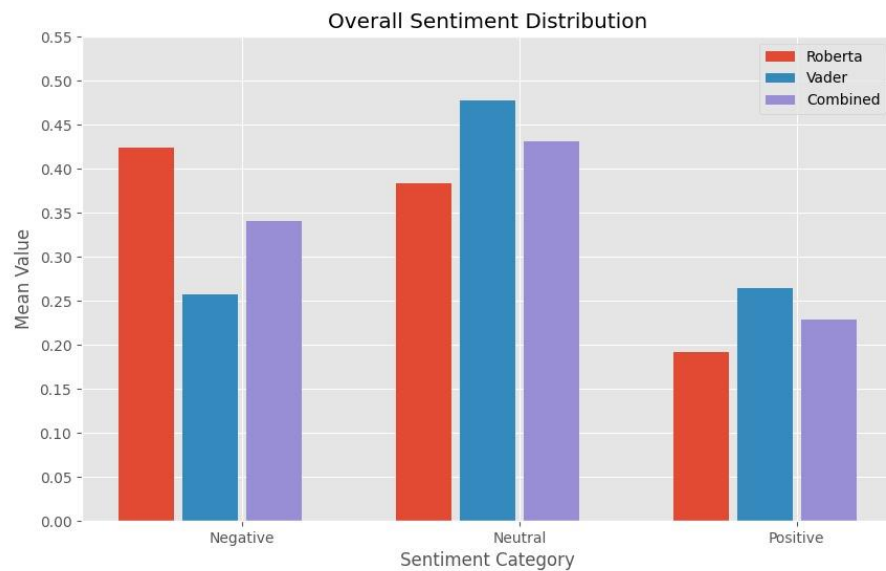Course Name & code: Deep Learning & 19CS7DETPL
Date: 1-12-2023

TITLE OF THE PROJECT	**Twitter Sentiment Analysis on the Israeli–Palestinian conflict using Vader Model**			
STUDENT NAME	Gauravdev HB	Fahad HS	Harsha S	Ashrith Reddy
USN	1DS20CS073	1DS20CS067	1DS20CS083	1DS20CS100
INDIVIDUAL CONTRIBUTION	Web Scraping The Data From Social Media Sites	Applying vader Models Required For Sentimental Analysis	Data Visualization using matplotlib and python	Applying Roberto Models Required For Sentimental Analysis
GUIDE	Prof Sarala			
PROJECT ABSTRACT :	This Twitter Sentiment Analysis project provides a nuanced exploration of public sentiment regarding the Israeli-Palestinian conflict by leveraging data from the Twitter platform. The study reveals a multifaceted spectrum of opinions and emotional responses, offering a comprehensive understanding of sentiment fluctuations during pivotal conflict events. The research underscores the influential role of social media, particularly Twitter, as a dynamic space for the expression and dissemination of diverse viewpoints. Despite unveiling valuable insights, the study acknowledges challenges inherent in sentiment analysis, including potential biases, intricate language nuances, and the risk of misinformation. The findings emphasize the necessity of continuous monitoring to capture evolving sentiment dynamics. Moreover, the project highlights the practical implications for diplomatic efforts and communication strategies, suggesting avenues for future research in navigating the complexities of public sentiment within the context of the Israeli-Palestinian conflict.			

PLATFORM USED (H/W & S/W TOOLS USED	Twitter, Pycharm, Python, JavaScript, Selenium, Node, Vader, react, Beautifull Soup
INTRODUCTION	Sentiment analysis, also known as opinion mining, is a powerful natural language processing (NLP) technique that involves the classification of text data to determine the emotional tone, attitude, or sentiment expressed within it. In the context of Twitter, sentiment analysis can provide valuable insights into public opinion, customer feedback, and trending topics. This analysis can help individuals, businesses, and organizations make data-driven decisions, understand their audience better, and gauge public sentiment towards various topics or products. Once Selenium has fetched the desired tweets, Beautiful Soup comes into play for extracting pertinent information from the HTML source code of the Twitter page. This involves parsing the HTML to extract critical elements such as tweet text, timestamps, usernames, retweets, likes, and other relevant metadata. Furthermore, Beautiful Soup can be employed to navigate through paginated search results or user timelines, ensuring data collection from multiple pages. Robust error-handling mechanisms should also be implemented to address any unforeseen issues that may arise during web scraping, such as missing elements or network disruptions. Visualizations play a vital role in comprehending the sentiment of the collected tweets. Creating histograms or pie charts can visually represent the distribution of sentiment categories in the dataset. Time series analysis can be used to plot sentiment trends over time, helping to identify shifts in public opinion. Word clouds, on the other hand, generate visual representations that highlight frequently occurring words or phrases associated with different sentiment categories.
LIST OF PAPERS/URLS REFERRED	1) Tweets are classify into the positive or negative comments using machine learning algorithm such as Naïve Bayes, Random forest (RF), Support vector machine (SVM), Unigram with Sentiwordnet and unigram with Sentiwordnet including negations are using as the input in this paper. Author derived three thousand one hundred eighty-four (3184) tweets using the tweeter API. Nine hundred fifty-four (954) positive, one thousand eighteen (1318) negative, 145 stop words have been identifying from 3184 tweets. Using. Author used feature of sentiment analysis like Bag of words (BOW), Term frequency vs Inverse document frequency (TF-IDF), Unigram with Sentiwordnet, Unigram with Sentiwordnet including negation words as an input. Author gets a conclusion that all the classifier with Unigram with Sentiwordnet and Unigram with Sentiwordnet including negation word shows higher accuracy the Bags of words (BOW) and term frequency vs Inverse document frequency (TF-IDF). Random forest algorithm with Unigram with Sentiwordnet including negation words get highest accuracy of 95.6%. 2) authors try to use machine learning algorithm for Arabic customer's feedback. They

	study two different type of methods which are voting and meta- classifier combination. They collecting data using Tweepy Application Programing Interface (API)17. There are many sarcastic and neutral tweets with the positive and negative tweets. Total 438,931 tweets were collected from that 75,774 are positive and 75,774 negative. Removing the al noisy data from the tweets like pictures, hashtags, retweets, emotions; second tokenization removing non Arabic letters, normalizing Arabic analogue letters. 10 classifiers NB, ME, LR, RR, PA, MNB, SVM, SGD and Ada boost BNB were used to extract and discover the polarity of given tweets. The highest accuracy achieved by PA and RR is 99.96%. Lowest accuracy achieved by Ada boost, LR and BNB which is less than 60%. 3) uses Amazon customer review data to find out the positivity, negativity and neutrality on customer's review. In this they compare two machine learning algorithms Naïve Bayes algorithm and Support vector machine (SVM). The input is the customer review of the Amazon products. The review maybe negative, positive or neutral. Apriori algorithm is used to extract the frequently used aspects from the input dataset. Sentiwordnet is used to calculate positivity, negativity and neutrality score and after that the classifier will apply. The comparison of the algorithm based on the performance can be calculated by using the Accuracy, Precision, Recall and F-1 Measure of each classification. By the experimental result Naïve Bayes classification is batter accuracy then Support vector machine (SVM). Calculation were done by True positive sample (TP), False positive samples (FP), True negative samples (TN) and False negative samples (FN).
DESIGN {SYSTEM DESIGN DIAGRAM}	Search and Navigation: Selenium will automate the process of entering search queries or visiting specific user profiles on Twitter. It can also handle the scrolling and loading of additional tweets as necessary. Data Extraction: Once the tweets are loaded, Beautiful Soup will be used to parse the HTML content and extract relevant tweet data. This includes the text of the tweets, timestamps, usernames, and any other information of interest. Preprocessing: Before conducting sentiment analysis, it's crucial to preprocess the tweet text. This may involve tasks such as text cleaning (removing special characters, emojis, and URLs), tokenization, and stemming or lemmatization. Sentiment Analysis: With the cleaned and preprocessed tweet data, sentiment analysis can be performed using various natural language processing (NLP) techniques or pre-trained models. Sentiment can be categorized as positive, negative, or neutral, providing insights into public sentiment on the given topic. Visualization and Reporting: Finally, the sentiment analysis results can be visualized using charts or graphs to provide a clear overview of the sentiment trends. This information can be used for reporting and decision-making purposes. By automating the tweet collection process with Selenium and extracting data using Beautiful Soup, we can efficiently gather large datasets for sentiment analysis, enabling

	us to gain valuable insights from Twitter's vast and dynamic conversation space. This approach is flexible and can be adapted for a wide range of research and analysis needs.
PROJECT SOURCE CODE LINK (GITHUB/ GOOGLE DRIVE)	www.github.com/Gauravdev456
CONCLUSION /FUTURE ENHANCEMENT	In conducting a Twitter Sentiment Analysis on the Israeli-Palestinian conflict, our study unveiled a rich tapestry of diverse perspectives and emotional responses within the Twitter community. Through the examination of user-generated content, we observed dynamic shifts in sentiment, often correlating with key events during the conflict. This analysis underscores the pivotal role of social media platforms, particularly Twitter, in shaping and reflecting public opinion in real-time. It also highlights the challenges of navigating biases, language nuances, and the potential for the spread of misinformation. Moving forward, our findings suggest the need for continuous monitoring and nuanced interpretation of social media sentiment to inform diplomatic efforts, communication strategies, and contribute to a more comprehensive understanding of public sentiment in this complex geopolitical landscape.

UI SCREENSHOTS OF SAMPLE RESULTS	

Neutral

REFERENCES	1. (Jul. 22, 2022). CDC. MPOX in the U.S. Centers for Disease Control Prevention. Accessed: Dec. 23, 2022. [Online]. Available: https://www.cdc.gov/poxvirus/monkeypox/about/index.html 2. A. Mandavilli. (Jul. 23, 2022). W.H.O. Declares Monkeypox Spread a Global Health Emergency. The New York Times. Accessed: Dec. 23, 2022. [Online]. Available: https://www.nytimes.com/2022/07/23/health/monkeypox-pandemic-who.html 3. (Aug. 4, 2022). A. S. P. for Affairs (ASPA). Biden-Harris Administration Bolsters Monkeypox Response; HHS Secretary Becerra Declares Public Health Emergency. Accessed: Dec. 23, 2022. [Online]. Available: https://www.hhs.gov/about/news/2022/08/04/biden-harris-administrationbolsters-monkeypox-response-hhs-secretary-becerra-declares-publichealth-emergency.html 4. J.-H. Yoo, ''Once bitten, twice shy: Our attitude towards Monkeypox,''J. Korean Med. Sci., vol. 37, no. 22, p. e188, 2022. 5. (Dec. 22, 2022). CDC. 2022 U.S. Map & Case Count. Accessed:Dec. 23, 2022. [Online]. Available: https://www.cdc.gov/poxvirus/monkeypox/response/2022/us-map.html