**Individual In-depth Report**

**Member name:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)

**Evaluated by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)

**Date: 10/09/2023**

**Tasks Assigned:**

* This report summarizes the paper “[Suspicious Activity Detection of Twitter and Facebook using Sentimental Analysis](https://drive.google.com/file/d/1N9bz3ZE4Wi4g64_UIoMzsDgOplt54iCA/view?usp=drive_link)”, which discusses the NLP techniques like Part of Speech (POS) tagging to get the syntactic meaning of the data. The authors have used the datasets from Twitter and Facebook to detect the suspicious behavior using Valance Aware Dictionary Sentiment Reasoner (VADER), a lexicon-based sentiment analysis.

**Summary:**

* In this study, two datasets were used: one with 5,000 tweets from 50 random Twitter accounts and another containing 12,000 posts from 50 randomly selected Facebook public accounts.
* Data preprocessing was crucial, serving as the initial step in Natural Language Processing (NLP) systems, particularly when handling unstructured data from social media. Data preprocessing involved several key steps. First, text cleaning removed noise like HTML, hashtags, and emoticons, followed by text normalization, which eliminated blanks, stop words, and small words. Text chunking, using Part-Of-Speech (POS) tagging, categorized words into nouns, verbs, or adjectives and segmented text strings into related phrases. Name Entity Recognition (NER) was used to classify text into predefined classes.
* The cleaned data was saved in .CSV format for sentiment analysis, employing Valance Aware Dictionary Sentiment Reasoner (VADER), a lexicon-based method. VADER mapped lexical features to emotion intensities, classifying sentiments as positive, negative, or neutral and identifying suspicious activities in Facebook and Twitter posts.
* VADER's scale ranged from -4.00 to +4.00, with 0 indicating neutrality. Multiple reviewers were used to ensure dictionary accuracy. Sentence-based features were normalized using Hutto Normalization with α = 15.
* In summary, this study emphasized data preprocessing's significance and the effectiveness of VADER for sentiment analysis in short social media texts. It facilitated sentiment classification and detected suspicious activities, making it a valuable tool for analyzing posts on platforms like Facebook and Twitter.

**Outcome:**

This study utilized NLP's Part-Of-Speech (POS) tagging to categorize words into nouns, verbs, and adjectives, aiding in text analysis. Additionally, Valence Aware Dictionary Sentiment Reasoner (VADER) proved valuable for sentiment analysis in social media data, providing a scalable and efficient approach to classify sentiments into positive, negative, and neutral categories, enhancing the identification of suspicious activities.

**References** *(with citation)*

[1] S. A. Mansoori, A. Almansoori, M. Alshamsi, S. A. Salloum, and K. Shaalan, “Suspicious Activity Detection of Twitter and Facebook using Sentimental Analysis,” TEM Journal, pp. 1313–1319, Nov. 2020, doi: 10.18421/tem94-01.

**Evaluation of Report**

**Evaluation summary with justification.**

The summary of the paper in this report is concise and clear. It explores NLP techniques like Part of Speech (POS) tagging to get the syntactic meaning of the data. This was used to perform sentiment analysis on social networking sites like Twitter and Facebook. The sentiment analysis technique used was Valance Aware Dictionary Sentiment Reasoner (VADER), a lexicon-based sentiment analysis.

**The quality of the major result(s) with justification.**  
  
The major results and outcomes of the reference paper have been discussed in detail in this report.

**The usefulness of the paper to the overall project.**   
  
The paper is highly relevant and contributes significantly to the overall project.

**Other comments**

No comments

**Evaluation Approval  
  
Evaluation by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu) **Date: 10/09/2023**

**Is the written report of the in-depth study complete with all the major result(s) of the paper(s)? If not, provide as many examples of the major result(s) missing in the written report as possible. (in bullet form). [Normally within 100 words]**

* Yes, all major results of the paper are covered.
* The in-depth study report is complete.

**Is each section of the guidelines sufficiently completed? If not, point out what is missing. [Normally within 40 words].**

Yes, each section of the guidelines is sufficiently completed.

**Is the quality of this version of the written report satisfactory? If not, then why not? [Normally within 40 words]**

Yes, the quality of this written report is satisfactory. The results of the paper contribute to achieving the goal of the overall project.

**Approval.  
  
Approved by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu) **Date: 10/09/2023  
  
Is the quality of this written in-depth study report and Evaluation report satisfactory? If not, then why not? (limit: 40 words)**

The quality of this written in-depth study report and Evaluation report is satisfactory. The evaluation report signifies correct evaluation and the report itself justifies the project topic.