**Individual Report**

**Member name:** [Avani Mundra](mailto:amudra@asu.edu)

**Evaluated by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)[Rahul Nayak](mailto:rrnayak@asu.edu)

**Date:** 09/03/2023

**Tasks Assigned:**

* Literature Review : Researched two specific research papers on data mining techniques to handle multimodal data
* Prepared questions related to information assurance and security in the project topic, thereby covering detailed subtopics with respect to suspicious activities in social media
* Reviewed the paper ‘Cybercrime Profiling : Text mining techniques to detect and predict criminal activities in microblog posts’ in detail and drafted an in-depth report for the same
* Evaluated [Yeshwanth Reddy Chennur](mailto:ychennur@asu.edu) ’s individual progress report and in-depth report for the present week

**Summary:**

* Researched on the subtopic: Data Mining techniques to handle multimodal data. The purpose of this subject is to find efficient data mining techniques that can help in analyzing and processing unstructured data which deals with dealing with suspicious activities on social media.
* In context of the above subtopic, reviewed the research paper ‘Cybercrime Profiling: Text mining techniques to detect and predict criminal activities in microblog posts’.
* The paper introduces the concept of cybercrime profiling, and discusses various data mining and text analytics techniques used in the past to investigate criminal activities in social media.
* The paper improves upon an already existing solution by implementing an approach to resolve the conflict of polysemy and synonymy by adding semantics for detecting suspicious profiles.
* The disambiguation step is added by introducing a similarity distance function that clusters and classifies the hashtags used in Twitter dataset into suspicious and non-suspicious posts.
* The Normalized Compression Distance function returns 0 when the text extracted from Twitter is similar to any word from the text corpus of suspicious words dataset.

**Outcome:**

The research paper improved the classification of text as suspicious by implementing a similarity distance function that introduces semantics and overcomes the challenges of synonymy and polysemy. It further enhances the system in terms of execution time.

**References:**

[16] S. Alami and O. Elbeqqali, “Cybercrime profiling: Text mining techniques to detect and predict criminal activities in microblog posts,” *International Conference on Intelligent Systems: Theories and Applications (SITA)*, Oct. 2015, doi: 10.1109/sita.2015.7358435.

**Evaluation of Report  
  
Evaluation by:** [Rahul Nayak](mailto:rrnayak@asu.edu) **Date:September 3, 2023**

**Is the weekly member report 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). [within 100 words]**

* Yes, The weekly member’s Report is complete in summarizing major results of the paper.
* It introduces the concept of cybercrime profiling and the summary of the paper has been written correctly in line with the topic chosen.

**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 the report is satisfactory.

**Approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)

**Date:** 09/03/2023