**Individual Report**

**Member name:** [Anuranjan Dubey](mailto:adubey37@asu.edu)

**Evaluated by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu)

**Date:** 09/08/23

**Tasks Assigned:**

* Literature review for Machine learning techniques applied in: “CyberTwitter: Using Twitter to generate alerts for Cybersecurity Threats and Vulnerabilities”
* Write In-depth Report while comparing it to methods seen before.
* Review and Evaluate 1 other person's report.

**Summary:**

* CyberTwitter collects real-time data from Twitter for cybersecurity intelligence.
* It uses a Security Vulnerability Concept Extractor (SVCE) to identify and extract cybersecurity-related entities. (Explained in-depth)
* The tool links extracted data to external semantic knowledge bases for comprehensive information.
* Intelligence is represented in RDF triples in a semantic knowledge base. (Explained in-depth)
* Analysts create system profiles with software and hardware information.
* The tool uses an intelligence ontology and SWRL rules (Explained in-depth) for reasoning.
* Analysts receive tailored threat alerts based on their system profiles.
* Alerts are time-sensitive, keeping analysts updated on emerging threats.
* CyberTwitter streamlines access to relevant real-time threat intelligence.
* The paper describes the CyberTwitter framework, which collects real-time data from Twitter using the Twitter Stream API and a set of keywords. The collected tweets are then cleaned using WordNet, a lexical database for English.
* The system was able to collect a dataset of 143,701 tweets, out of which 10,004 relevant tweets were used to create intelligence entities. Based on an organization's system profile, the system issued 15 threat alerts related to software vulnerabilities .
* The paper concludes, including the use of semantic textual similarity systems to improve tweet selection, recognizing relevant tweets based on their content, and expanding the word embedding model trained on general text to include cybersecurity text.

**Outcome:**

The CyberTwitter system's performance was assessed through a ten-day experiment, with most alerts proving valuable. It effectively extracts cybersecurity insights from real-time Twitter data, demonstrating its capacity for issuing threat alerts.

**References** *(with citation)*

[1] S. Mittal, P. K. Das, V. Mulwad, A. Joshi and T. Finin, "CyberTwitter: Using Twitter to generate alerts for cybersecurity threats and vulnerabilities," *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, USA, 2016, pp. 860-867, doi: 10.1109/ASONAM.2016.7752338.

[2] N. Dionísio, F. Alves, P. M. Ferreira and A. Bessani, "Cyberthreat Detection from Twitter using Deep Neural Networks," *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8, doi: 10.1109/IJCNN.2019.8852475.

**Evaluation of Report  
  
Evaluation by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu) **Date:** 09/10/23

**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]**

* The weekly report is complete with the major result, which is to extract cybersecurity intelligence and issue alerts for any threats generated, emphasized.
* In-depth analysis of each of the steps required to find the final result is completed successfully.

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

With all the points covered in detail, 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 written report has touched upon all the major points in the research paper and hence is of good quality.

**Approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date: 09/11/2023**