**Individual Report**

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

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

**Date:** 10/21/23

**Tasks Assigned:**

* Literature review for time series based machine learning techniques applied in: “A Framework to Predict Social Crime through Twitter Tweets By Using Machine Learning”
* Read others' not so Important Papers.
* Review and Evaluate [Rahul Nayak](mailto:rrnayak@asu.edu)’s individual progress report.

**Summary:**

* Searched for new papers with similar or different approaches, read others' not so Important Papers, evaluated Rahul’s Report, compared the paper with time series aspects.
* The purpose of the research paper's proposed framework is to predict major types of social media crimes using Twitter data.
* The framework comprises three modules: data pre-processing, a classifier model builder, and prediction.
* Machine learning algorithms used in the framework include Multinomial Naive Bayes, K-Nearest Neighbors, and Support Vector Machine.
* The framework classifies data into crime categories, including Cyber-Stalking, Cyber-Harassment, Cyber-Bullying, Cyber-Scam, and Cyber-Hacking.
* It achieves higher accuracy compared to existing network-based feature selection approaches, with an accuracy of 92% using Bag-of-Words.
* The SKRAM model is used to identify hacking behavior in social media users, considering factors like skills, knowledge, resources, access to targets, and motivation.
* The proposed methodology involves data collection, pre-processing, and classification using machine learning, with N-Gram language models for accuracy measurement.
* The document discusses the significance of social media data for various investigations and predicts future trends.
* The proposed framework can be useful for law enforcement agencies in reducing crime rates and protecting society.
* Additional crime categories, real-time Twitter data streaming, and system enhancements are suggested for future work.

**Outcome:**

* While the paper showed significant results and an excellent accuracy, it might not be feasible for real time or temporal data.
* Why we marked this as not so important; We already have focused on traditional machine learning and data mining approaches with similar static data which offer advantages but are limited for real time detection or time series data.

**References** *(with citation)*

[3] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante and K. Veeramachaneni, "TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks," 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 33-43, doi: 10.1109/BigData50022.2020.9378139.

[10] Z. Abbass, Z. Ali, M. Ali, B. Akbar and A. Saleem, "A Framework to Predict Social Crime through Twitter Tweets By Using Machine Learning," 2020 IEEE 14th International Conference on Semantic Computing (ICSC), San Diego, CA, USA, 2020, pp. 363-368, doi: 10.1109/ICSC.2020.00073.

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

**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]**

Yes, the weekly report is complete with all the important results of the paper explained in detail. The machine learning algorithms used for detecting social media crimes on Twitter are accurately described and the one with the highest accuracy value is found.

**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 detailed and complete.

**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 written report is satisfactory.

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

**Date: 10/23/2023**