**Individual Report**

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

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

**Date:** 09/22/23

**Tasks Assigned:**

* Literature review for Machine learning techniques applied in: “An effective security alert mechanism for real-time phishing tweet detection on Twitter”.
* Swapped an Irrelevant paper [2] to the important one [1].
* Write In-depth Report while comparing it to methods seen before.
* Review and Evaluate Rahul’s Report.
* Preparing Midterm Report
* Preparing Individual Midterm Review Report

**Summary:**

* Prepared midterm report and individual review report, swapped and identified an irrelevant paper, evaluated Rahul’s Individual report, wrote in depth report relevant to the project topic and machine learning methods.
* Research Paper: Phishing attacks have extended to social media platforms like Twitter due to their popularity and difficulty in detection.
* Challenges in detecting phishing on Twitter include rapid information dissemination, detection difficulty, lack of reciprocity, real-time nature, and limited content space.
* Feature extraction involved 22 features from various categories, combined with 7 additional features from the WEKA tool.
* WEKA is a machine learning tool used for feature selection and evaluation.
* The security alert mechanism has two stages: machine learning training (Random Forest) and mechanism formulation for real-time detection.
* Machine learning training used a dataset with 2973 training data containing URLs from Twitter, labeled as phishing or safe.
* The machine learning methods used and compared primarily focus on the Random Forest algorithm.
* The security alert mechanism mainly detects phishing URLs.
* A proposed classification model using Random Forest achieved high accuracy in phishing tweet detection (94.64% precision, 95.49% recall).

**Outcome:**

The Random Forest-based security alert mechanism demonstrated impressive performance with a precision of 94.64% and a recall of 95.49% in phishing tweet detection through simulation.

When implemented and evaluated with real Twitter and PhishTank datasets, the mechanism achieved an impressive 97.50% accuracy in promptly alerting users to phishing URLs.

**References** *(with citation)*

[1] Seow Wooi Liew, Nor, Mohd Taufik Abdullah, Razali Yaakob, and Mohd Yunus Sharum, “An effective security alert mechanism for real-time phishing tweet detection on Twitter,” Comput. Secur., vol. 83, pp. 201–207, 2019, doi:<https://doi.org/10.1016/J.COSE.2019.02.004>.

[2] 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.

**Evaluation of Report  
  
Evaluation by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu) **Date: 09/24/23**

**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 member report is complete with all the major results of the paper mentioned. The machine learning algorithm mainly used in the paper, Random Forest algorithm has been evaluated and the values of its accuracy, precision and recall have been calculated. It has been successfully found from the paper that these parameters have high values.

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

Yes, each section of the report is sufficiently detailed and aligns with the guidelines.

**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: 09/24/2023**