**Individual Report**

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

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

**Date:** 29th September 2023

**Tasks Assigned:**

* Literature review for Machine learning techniques applied in: “A Feature Based Approach to Detect Fake Profiles in Twitter”.
* Identifying previously read papers where methods employed in these papers like feature selection, extraction with machine learning approaches were similar.
* Write In-depth Report.
* Review and Evaluate Rahul’s Report.
* Searched for Additional Papers

**Summary:**

* Searched additional research papers, Identified the similar methodologies, wrote an in-depth report and reviewed Rahul’s Indepth report and Individual report.
* The feature-based approach for spotting fake Twitter accounts utilized a set of 24 distinct features.
* In the study, three machine learning algorithms were employed and compared: Logistic Regression, Support Vector Machines (SVM), and Random Forest.
* Among these algorithms, Random Forest exhibited the highest performance based on accuracy, precision, recall, and F-score metrics.
* It's not explicitly stated whether this approach can be directly applied to social media platforms other than Twitter and Facebook. The study primarily focuses on these platforms.
* Several data mining tools were mentioned like snippets, including DeepScan, SybilBlind, Botometer, and COLOR+.
* The proposed methodology involves data preprocessing, feature selection, model training with supervised machine learning algorithms, and evaluation using a test dataset. The dataset is split into an 80:20 ratio for training and testing purposes.
* The primary objective is to offer a user-end solution for combating fake accounts and automated bots on social media platforms.
* The study doesn't explicitly state its applicability to social media platforms beyond those mentioned, leaving uncertainty about its generalizability.

**Outcome:**

The evaluation found Random Forest with 97.9% accuracy outperformed Logistic Regression (95.7%) and SVM (80.8%) in detecting fake social media accounts.

Future work includes studying human-operated fake accounts, adding features for better bot detection, incorporating tweet sentiment analysis, and developing a real-time web browser-based tool to identify bot accounts.

**References** *(with citation)*

[1] J. Kaubiyal and A. K. Jain, “A feature based approach to detect fake profiles in twitter,” Melbourn, VIC, Australia: Association for Computing Machinery, 2019, pp. 135–139. doi: https://doi.org/10.1145/3361758.3361784. Available: https://doi.org/10.1145/3361758.3361784

**Evaluation of Report  
  
Evaluation by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu) **Date: 1st October 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 report is complete. The report covers all the tasks done by the reporter during the course of the week ending 2nd October 2023. Each task has been explained in detail.

**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 covers the important tasks as well as a summary of the research paper studied by the reporter.

**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 version of the written report is satisfactory. It does justification to the tasks completed by the reporter this week along with giving a brief summary of the research paper chosen by the reporter for study this week.

**Approved by:**[Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date:10/02/2023**