**Individual Report**

**Member name:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu)

**Evaluated by:** [Justin Young](mailto:jtyoun15@asu.edu) [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)

**Date:** 09/29/23

**Tasks Assigned:**

* Preparing Weekly Report along with [Justin Young](mailto:jtyoun15@asu.edu)
* Preparing and Evaluating Individual Progress Report
* Preparing and Evaluating Individual In-depth Report
* Performing the literature review of research papers assigned.
* Preparation of an in-depth report of one research paper among the research papers found – ‘Ecosystem of Spamming on Twitter: Analysis of Spam Reporters and Spam Reportees’.

**Summary:**

* All the tasks assigned for this week have been completed successfully.
* Reviewed the research paper – ‘Ecosystem of Spamming on Twitter: Analysis of Spam Reporters and Spam Reportees’.
* Analysis of spam activity and identification of reporters and reportees on Twitter is done.
* Incorrect classification of honest users as spam is avoided in the algorithm proposed in the paper.
* The algorithm proposed has the following main steps –
  + Data Collection – A dataset containing tweets is obtained from which reporters and reportees are separated. This is done using Python as a programming language.
  + Categorization and Feature Set Description – This step involves identification of different classes of reporters and reportees. Post identification, classification of reporters and reportees is done using metrics or attributes, also known as feature sets.
  + Applying Data Mining Algorithms – Decision Tree, K-Nearest Neighbors and Random Forest Classifier are the main algorithms applied for data mining. This step is done using R as a programming language.
* The best algorithm is chosen based on the accuracy. The accuracy is measured and verified using manual classification.
* The final result obtained from the research paper is that the decision tree algorithm is the best classifier for reporters and reportees.

**Outcome:**

Tasks assigned for the week have been completed successfully. From the research paper, it is found that for the process of classification of reporter and reportee, the Decision Tree algorithm is the best due to its high accuracy value.

**References** *(with citation)*

[33] P. Sinha, O. Maini, G. Malik and R. Kaushal, "Ecosystem of spamming on Twitter: Analysis of spam reporters and spam reportees," 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 2016, pp. 1705-1710, doi: 10.1109/ICACCI.2016.7732293.

**Evaluation of Report  
  
Evaluation by:** [Justin Young](mailto:jtyoun15@asu.edu) **Date: 09/30/23**

**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, all major results of the chosen research paper were covered.

**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 have been 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 this version of the written report is satisfactory. All the tasks have been completed and the summary of the research paper is also complete.

**Approved by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu) **Date:** 09/30/2023