**Individual Report**

**Member name:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)

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

**Date: 10/23/2023**

**Tasks Assigned:**

* Preparing Weekly Report
* Preparing [Work\_Distribution\_Draft](https://docs.google.com/document/d/1DUjMSRDReQVWcPMbvv48WAdtmWlTxwvt/edit?usp=drive_link&ouid=107121383631492907164&rtpof=true&sd=true)
* Preparing individual progress report
* Added summary of not so important reference paper [4] in the summary document: [Summary\_Statement\_Not\_So\_Important\_Papers](https://docs.google.com/document/d/1QmxaOqG8rONc0v80Qa2wFGUMXmtyd_5nVvCeJUYJvuM/edit?usp=share_link)
* Read 5 papers from [The "Not-So-Important" Reference Papers](https://drive.google.com/drive/folders/1MpdS68DsMj4vHUt9yutzkOcZJDGQ6NkM?usp=share_link) from the References List.
* Evaluating and approving team members’ individual progress
* Evaluating and approving GANTT Chart
* Assigning the task distribution
* Providing a rough draft of content for Weekly report and GANTT Chart
* Organizing group meetings
* Organizing google drive

**Summary:**

* All of the tasks were successfully finished. Creating an individual progress report, creating a weekly report, work distribution plan, and adding a summary of one not so important paper in a summary document linked above, were among my responsibilities. Furthermore, I evaluated and obtained approval for the individual progress reports that 4 team members provided. The GANTT Chart was also reviewed and approved. Task assignments were distributed, along with GANTT charts and weekly report summaries as needed. Additionally, held group meetings and arranged Google Drive files in accordance with the requirements.
* Summarized not so important paper titled “[A Multiple Feature Category Data Mining and Machine Learning Approach to Characterize and Detect Health Misinformation on Social Media](https://drive.google.com/file/d/1prOq20Wg0ybLtr4LGisbEDf-DdpOjasE/view?usp=drive_link)”. This article examines the spread of false information about health on social media, with a particular emphasis on the 2016 Twitter debate on Zika. After identifying popular tweets that contain both false and accurate information, the study creates a data mining strategy to extract attributes for the purpose of detecting disinformation. It takes into account signal-based features, content and language features, user features, and information dissemination features. These features are used to train classifiers such as random forest (RF) and support vector machine (SVM). When all feature categories were put together, the RF classifier detected disinformation with over 83% accuracy and over 90% AUC. Responding to false information during health emergencies such as the COVID-19 pandemic and the Zika epidemic can benefit greatly from this strategy.

**Outcome:**

I successfully completed tasks, evaluated team reports, and reviewed GANTT Charts. The study examined health misinformation on social media, focusing on the 2016 Zika debate. It used data mining for detection, considering signal, content, user, and information dissemination features. The random forest classifier achieved over 83% accuracy and 90% AUC. This approach is valuable during health crises like Zika and COVID-19.

**References** *(with citation)*  
  
[1] L. Safarnejad, Q. Xu, Y. Ge and S. Chen, "A Multiple Feature Category Data Mining and Machine Learning Approach to Characterize and Detect Health Misinformation on Social Media," in IEEE Internet Computing, vol. 25, no. 5, pp. 43-51, 1 Sept.-Oct. 2021, doi: 10.1109/MIC.2021.3063257.

**Evaluation of Report  
  
Evaluation by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu) **Date: 10/23/2023**

**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 member report is complete with all the major results of the paper. All the tasks assigned to the member have been completed within the given time frame. The summary of the research papers studied, explores the use of data mining for detection, considering signal, content, user, and information dissemination features. These results of the paper prove that the data mining techniques are really accurate and can contribute to the overall goal of our project.

**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 are 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. The summary of the “Not so Important” papers written in this report is concise and clear.

**Approved by:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu) **Date: 10/23/2023**