**Individual In-depth Report**

**Member name:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)

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

**Date: 09/23/2023**

**Tasks Assigned:**

* Focus on my question to work upon as my area of research for this project.
* The question I chose was to do my research and find a system that maintains the balance between detecting suspicious activity in social media and safeguarding the user's privacy and maintaining data integrity.
* Review of [**Comprehensive Survey on Big Data Privacy Protection**](https://drive.google.com/file/d/1lnxy_kihxdIVJHpah0283JycADq59Jwx/view?usp=drive_link)

**Summary:**

* The main focus area of this paper is to elaborate on the existing challenges and unresolved issues in privacy-preserving data mining (PPDM).
* A key issue of PPDM is how to manipulate data using a specific approach to enable the development of a good data mining model on modified data, thereby meeting a specified privacy need with minimum loss of information for the intended data analysis task.
* The current review study aims to utilize the tasks of data mining operations without risking the security of individuals’ sensitive information, particularly at the record level.
* In this survey, the privacy preservation in the big data life cycle is considered at the data preprocessing and data mining task stages.
* PRIVACY PRESERVATION IN DATA PROVIDERS:
  + Data providers cannot disclose any information because they regard their data as extremely sensitive.
  + Data providers opt to never release private information because they are aware of the value of their data to data collectors.
  + Data providers may be willing to disclose some of their sensitive information for specific rewards, such as improved services or financial benefits.
  + If data providers cannot either block access to their sensitive information, then data collectors can misrepresent data collected by the data providers in such a manner that actual information cannot be easily revealed.
* PRIVACY PRESERVATION IN DATA COLLECTION:
  + The goal of preserving data privacy is to safeguard privacy during data collection and transmission to different data mining servers by finding the minimum portion of private information required to construct accurate data mining models
  + The direct disclosure of data to data miners will infringe on the privacy of data providers, particularly in cases where data miners execute mining algorithms using the data provided by data collectors and extract valuable information from data.
  + Thus, three types of approaches have been generally developed to conceal the raw data from their original value - data exchange, data cryptographic, and data modification.
* The other approaches to privacy preserving data mining techniques include data modification approaches like k-anonymity approach, l-diversity approach, and t-closeness approach.

**Outcome:**

* The existing PPDM techniques are intensively reviewed and classified based upon their methods that used data modification approaches, which then represented the main contribution of this study that will help researchers in this field having comprehensive understanding of PPDM.
* PPDM continues to have potential challenges and open issues that would open the door for further research in the future.

**References** *(with citation)*  
  
[6] Mohammed Binjubeir, Abdulghani Ali Ahmed, Mohd Arfian Bin Ismail, Ali Safaa Sadiq, Muhammad Khurram Khan, “Comprehensive Survey on Big Data Privacy Protection,” in “IEEE Access, Vol 12”, Riyadh, Saudi Arabia, 2019, pp.20067-20079

**Evaluation of Report**

**Evaluation summary with justification.**

This report explores privacy-preserving data mining (PPDM) to enhance social media surveillance, striking a balance between user privacy and data integrity. It addresses data provider concerns, utilizing methods like data modification and encryption. This research offers insights for safeguarding user privacy while detecting suspicious activities on social media.

**The quality of the major result(s) with justification.**  
This study's major result is effective privacy-preserving data mining techniques, ensuring social media surveillance while safeguarding user privacy and data integrity.

**The usefulness of the paper to the overall project.**   
The paper provides essential insights into privacy-preserving data mining for robust social media surveillance while safeguarding privacy and data integrity.

**Other comments**

None

**Evaluation Approval  
  
Evaluation by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date: 09/23/2023**

**Is the written report of the in-depth study 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). [Normally within 100 words]**Yes, the written report of the in-depth study is complete with all the major results of the paper. This report critically examines privacy-preserving data mining (PPDM), emphasizing its significance in detecting suspicious activity on social media platforms while safeguarding user privacy and data integrity. It highlights the challenges in manipulating data to strike a balance between privacy and information loss. The study addresses data provider concerns and explores methods like data modification, cryptographic techniques, and anonymization to ensure privacy throughout the data lifecycle. This research contributes valuable insights for robust social media surveillance while upholding data integrity and user privacy.

**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 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.

**Approval.  
  
Approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date: 09/24/2023  
  
Is the quality of this written in-depth study report and Evaluation report satisfactory? If not, then why not? (limit: 40 words)**

Yes, the report is evaluated by me and the report has been studied carefully with all the remarks added in the evaluation report. The quality of this written in-depth study report and evaluation report is satisfactory.