**Individual In-depth Report**

**Member name:** [Avani Mundra](mailto:amudra@asu.edu)

**Evaluated by:** [Rahul Nayak](mailto:rrnayak@asu.edu)

**Date: 10/02/2023**

**Tasks Assigned:**

* This report summarizes the paper ‘Social Big Data Mining Framework for Extremist Content Detection in Social Networks’ in depth.

**Summary:**

* The research paper focuses on leveraging social media, to analyze user-generated content with a particular emphasis on detecting extremist and violent content.
* The proposed framework aims to automate the process of collecting, analyzing, and monitoring public text posts on Facebook using natural language processing (NLP) techniques and big data analytics.
* The framework consists of several key phases, including data extraction from Facebook through API graph calls, sentiment analysis, extremist content detection, storage in HBase (a distributed database), and the development of a monitoring interface for law enforcement personnel.
* The paper discusses the challenges of sentiment analysis in social media due to the contextual nuances of opinion words and the informal language used in platforms like Facebook.
* Various NLP techniques, such as translation services, website mapping, authorship attribution, and linguistic identifiers, are explored as tools for content analysis and identification of radical behaviors.
* The integration of sentiment analysis with behavioral linguistic identifiers is proposed to enhance the detection of extremist behavior, making it more effective in identifying potential threats.
* The paper underscores the potential of big data analytics and distributed processing techniques for improving the efficiency and scalability of NLP algorithms in monitoring and analyzing social media data, particularly for early detection of lone-wolf terrorists and enhancing public safety through advanced surveillance tools.

**Outcome:**

The research paper presents a comprehensive framework that combines natural language processing, sentiment analysis, and behavioral linguistic identifiers within a big data analytics context to automate the detection of extremist and violent content in public text posts on Facebook. This framework has the potential to significantly enhance law enforcement efforts in monitoring social media for early detection of lone-wolf terrorists and improving public safety through advanced surveillance and analysis tools.

**References**

[1] E. Mouhssine and C. Khalid, "Social Big Data Mining Framework for Extremist Content Detection in Social Networks," 2018 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Rabat, Morocco, 2018, pp. 1-5, doi: 10.1109/ISAECT.2018.8618726.

**Evaluation of Report**

**Evaluation summary with justification.**

The research paper offers a relevant framework to detect extremist content on Facebook via NLP and big data analytics. It acknowledges challenges like informal language and proposes integrative NLP methods. While emphasizing public safety and law enforcement, ethical considerations and empirical validation could strengthen its impact and credibility.

**The quality of the major result(s) with justification.**

The paper's major result highlights the integration of sentiment analysis with linguistic identifiers, enhancing the detection of extremist behavior. Additionally, it emphasizes the potential of big data analytics and distributed processing, offering more efficient NLP algorithms for early detection of lone-wolf terrorists and advancing public safety through advanced surveillance tools.

**The usefulness of the paper to the overall project.**   
The NLP techniques used in this paper can be further used to detect malicious activities on social media, successfully targeting the main goal of this project.

**Other comments**

No Comments

**Evaluation Approval  
  
Evaluation by:** [Rahul Nayak](mailto:rrnayak@asu.edu) **Date: 10/02/2023**

**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 weekly member report is considered complete as it successfully includes all the major findings and insights from the paper. The provided text offers a comprehensive summary of the research paper, encompassing its objectives, the application of data mining techniques, sentiment analysis, and the proposed framework for detecting extremist content in social networks.

**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 report is satisfactory.

**Approval.  
  
Approved by:**[Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date:10/02/2023  
  
Is the quality of this written in-depth study report and Evaluation report satisfactory? If not, then why not? (limit: 40 words)**

The quality of this written in-depth study report and Evaluation report is satisfactory. The evaluation report signifies correct evaluation and the report itself justifies the project topic by addressing an application of data mining techniques, and sentiment analysis.