## **Group 1-5 Project Plan**

**Project Title**: Detection Process of Suspicious Activities on Social Media Using Data Mining and Machine Learning.

## 

## List of members

**Group Leader:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)

**Group Deputy Leader:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)

**Group Members:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)

* [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)
* [Avani Mundra](mailto:amudra@asu.edu)
* [Justin Young](mailto:jtyoun15@asu.edu)
* [Anuranjan Dubey](mailto:adubey37@asu.edu)
* [Rahul Nayak](mailto:rrnayak@asu.edu)
* [Sangeeth Santhosh](mailto:ssantho9@asu.edu)
* [Yeshwanth Reddy Chennur](mailto:ychennur@asu.edu)

## 

## Project Scope

The project will focus on the following areas:

* **Data Exploration:**
  + There are a wide variety of suspicious activities that go on social media these days. Thus it is important that we study the different types of suspicious activities. [1] [2] [3] [4]
* **Secure Data Collection Techniques:**
  + Finding the right balance between monitoring or detecting suspicious activity or threats in Social media and respecting the user's privacy to maintain ethics, data confidentiality and integrity is crucial while dealing with real data. [5] [6] [7] [8] [9]
* **Data Preprocessing and Feature Engineering:**
  + Discussing the various preprocessing and/or NLP techniques to produce a robust dataset for detection of suspicious activities on social media. [10] [11] [12] [13] [14] [38]
* **Data Mining Techniques to Handle Multimodal Data:**
  + The data obtained from Social Media can be unstructured and varied as it can contain texts, images, videos and more. So we need to find a data mining technique that can help us process and analyze suspicious activity. [15] [16] [17] [18] [19]
* **Leveraging Machine Learning Models:**
  + Specific machine learning algorithms have been previously employed to enhance the accuracy and efficiency of the proposed tools in the previous research in classifying and categorizing security-related content from Twitter. Research will be performed on whether these algorithms be extended beyond their initial implementation on Twitter, and if they can be adapted for use on other social media platforms with minimal modifications. [20] [21] [22] [23] [37]
* **Model Security Assurance:**
  + Discussion on techniques producing a secure machine learning detection model which can refrain from tampering or unauthorized modifications from hackers. [24] [25] [26] [27] [28] [29] [30]
* **Model Evaluation:**
  + The suspicious social media accounts to be analyzed need to be found correctly without incorrectly grouping non-suspicious accounts as suspicious. Such incorrect classification would mean questioning the integrity of an honest user. Research will be performed on whether the right approach can be found to classify social media accounts into suspicious or non-suspicious without false grouping. [31] [32] [33] [34] [35]

## Project Outcomes:

The project is expected to achieve the following outcomes:

* **Machine Learning Models:** 
  + Finding reliable machine learning models that can recognize different kinds of suspicious activity on social media networks. The suspicious activities can be fraudulent accounts, hate speech, cyberbullying, and misinformation, multiple models may be created.
* **Accurate Results :**
  + Coming up with the research of the machine learning models which produces high levels of accuracy, precision, recall, and F1-score when detecting instances of suspicious activity.
* **Efficient Data Processing:**
  + Focus on finding efficient techniques to process large volumes of data without compromising the performance rate at any instance.
  + Conduct an in-depth analysis to determine the suspicious activities which would be identified using specific features and indicators.
* **Adaptive Learning :**
  + A system that’s designed to adapt and evolve to new forms of threats and suspicious activities.
  + The machine learning models would not only make accurate predictions but also provide insights into why a particular activity is deemed suspicious.
* **Privacy-Preserving Techniques :**
  + A system that protects the integrity and confidentiality of user data while detecting social media for suspicious activities and threats.
  + Maintains the ethical standards of user data and ensures data privacy.
* **Documentation and Reporting Findings of the Project:** 
  + A thorough report describing the project's methodology, algorithms, steps of implementation, and outcomes will be prepared.
* **Basis for Further Research:** 
  + The models and conclusions will operate as a starting point for additional study in the area of detecting suspicious behavior on social media.

## Project Timeline

The following is an Initial Gantt chart that shows the timeline for the project:

**Link to the Gantt chart:** [Gantt Chart](https://docs.google.com/spreadsheets/d/1TpkQiW_teJU5ORx2j8N8k_ihbj1cB_Zh/edit#gid=1653990051)

**Gantt chart updated by**: [Krupaben Kothadia](mailto:kkothadi@asu.edu)[Justin Young](mailto:jtyoun15@asu.edu)

**Gantt chart approved by**: [Krupaben Kothadia](mailto:kkothadi@asu.edu)**,** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)

## Project Risks (Optional)

The following are some of the risks associated with the project:

* **Incomplete or Inaccurate Information:**
  + Information on different data mining techniques, machine learning models, or security measures may be insufficient or inaccurate in the literature.
* **Dependency on External Sources:**
  + The literature study depends on the accessibility and dependability of outside sources, which may be modified or stopped altogether.
* **Complexity of Model Evaluation:**
  + It can be difficult to evaluate machine learning models and categorize social media accounts, and there may not be a single set of accepted standards.
* **Security Risks:**
  + When discussing machine learning model security, it's possible to unintentionally reveal information that could be used by malicious individuals.
* **Zero-Day Detection Challenges:**
  + Zero-day detection techniques might not always work or have practical application restrictions.
* **Data Mining Challenges:**
  + It may be difficult to find appropriate data mining methods to handle the variety of unstructured social media data, and some methods might be less efficient than others.
* **Resource Constraints:**
  + The quality and breadth of the literature review may suffer from a lack of time and resources.

## References [Ref number] Author’s initials. Author’s Surname, “Title of paper,” in “Name of Conference”, Location, Year, pp. Xxx

## [1] Xianghan Zheng, Zhipeng Zeng, Zheyi Chen, Yuanlong Yu, Chunming Rong,

## “Detecting spammers on social networks”,Neurocomputing,Volume 159,2015,

## Pages 27-34,ISSN 0925-2312

[2] S. Gheewala and R. Patel, "Machine Learning Based Twitter Spam Account Detection: A Review," *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2018, pp. 79-84, doi: 10.1109/ICCMC.2018.8487992.

[3] S. Shinde and S. B. Mane, "Malicious Profile Detection on Social Media: A Survey Paper," *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 2021, pp. 1-5.

[4] M. Jiang, P. Cui and C. Faloutsos, "Suspicious Behavior Detection: Current Trends and Future Directions," in *IEEE Intelligent Systems*, vol. 31, no. 1, pp. 31-39

[5] Xinjun Qi, Mingkui Zong, “An Overview of Privacy Preserving Data Mining,” in *“International Conference on Environmental Science and Engineering (ICESE)”*, Harbin, China, 2011, pp.1341-1347

[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

[7] J. Vaidya, C. Clifton, “Privacy-preserving data mining: why, how, and when”, in *“IEEE Security & Privacy, Vol 2, Issue 6”*, USA, 2004, pp.19-27

[8] M. S. Sudha, K. A. Priya, A. K. Lakshmi, A. Kruthika, D. L. Priya and K. Valarmathi, "Data Mining Approach for Anomaly Detection in Social Network Analysis," *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India, 2018, pp. 1862-1866, doi: 10.1109/ICICCT.2018.8472985.

[9] Anurag P. Jain, Vijay D. Katkar, “Sentiments analysis of Twitter data using data mining,” in *“International Conference on Information Processing (ICIP)”*, Pune, India, 2015, pp.807-810

[10] S. Gharge and M. K. Chavan, “An integrated approach for malicious tweets detection using NLP,” *International Conference on Inventive Communication and Computational Technologies,* Mar. 2017, doi: 10.1109/icicct.2017.7975235.

[11] A. Kumar and S. Singh, “Detection of User Cluster with Suspicious Activity in Online Social Networking Sites,” *2013 Second International Conference on Advanced Computing, Networking and Security*, Dec. 2013, doi: 10.1109/adcons.2013.17.

[12] M. S. Al-Ghamdi and M. A. Khan, “Intelligent analysis of Arabic tweets for detection of suspicious messages,” *Arabian Journal for Science and Engineering*, vol. 45, no. 8, pp. 6021–6032, Mar. 2020, doi: 10.1007/s13369-020-04447-0.

[13] M. P. Selvan and R. Selvaraj, “Mining user Message Pattern for Suspicious Behavior on Terrorism using NLP in Social Networks with Single Sign-On,” *Indian Journal of Science and Technology*, Apr. 2017, doi: 10.17485/ijst/2017/v10i14/111364.

[14] S. A. Mansoori, A. Almansoori, M. Alshamsi, S. A. Salloum, and K. Shaalan, “Suspicious Activity Detection of Twitter and Facebook using Sentimental Analysis,” *TEM Journal*, pp. 1313–1319, Nov. 2020, doi: 10.18421/tem94-01.

[15] S. Ghosal, A. Jain and D. K. Tayal, "An approach to detect abusive content incorporating Word2Vec and Multilayer Perceptron," *2022 IEEE Bombay Section Signature Conference (IBSSC)*, Mumbai, India, 2022, pp. 1-5, doi: 10.1109/IBSSC56953.2022.10037274.

[16] S. Alami and O. Elbeqqali, “Cybercrime profiling: Text mining techniques to detect and predict criminal activities in microblog posts,” *International Conference on Intelligent Systems: Theories and Applications (SITA)*, Oct. 2015, doi: 10.1109/sita.2015.7358435.

[17] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin, “TargetVue: Visual Analysis of Anomalous user Behaviors in Online Communication systems,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 280–289, Jan. 2016, doi: 10.1109/tvcg.2015.2467196.

[18] Ravneet Kaur, Sarbjeet Singh. "A survey of data mining and social network analysis based anomaly detection techniques." *Egyptian Informatics Journal*, Volume 17, Issue 2, 2016, Pages 199-216. ISSN 1110-8665. doi: 10.1016/j.eij.2015.11.004.

[19] S. Garg, K. Kaur, N. Kumar and J. J. P. C. Rodrigues, "Hybrid Deep-Learning-Based Anomaly Detection Scheme for Suspicious Flow Detection in SDN: A Social Multimedia Perspective," in *IEEE Transactions on Multimedia*, vol. 21, no. 3, pp. 566-578, March 2019, doi: 10.1109/TMM.2019.2893549.

[20] S. Mittal, P. K. Das, V. Mulwad, A. Joshi and T. Finin, "CyberTwitter: Using Twitter to generate alerts for cybersecurity threats and vulnerabilities," *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, USA, 2016, pp. 860-867, doi: 10.1109/ASONAM.2016.7752338.

[21] A. Ritter, E. Wright, W. Casey, and T. Mitchell, “Weakly supervised extraction of computer security events from twitter,” Florence, *Italy: International World Wide Web Conferences Steering Committee*, 2015, pp. 896–905. doi: https://doi.org/10.1145/2736277.2741083.

[22] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante and K. Veeramachaneni, "TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks," *2020 IEEE International Conference on Big Data (Big Data)*, Atlanta, GA, USA, 2020, pp. 33-43, doi: 10.1109/BigData50022.2020.9378139.

[23] N. Dionísio, F. Alves, P. M. Ferreira and A. Bessani, "Cyberthreat Detection from Twitter using Deep Neural Networks," *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8, doi: 10.1109/IJCNN.2019.8852475.

[24] G. Rathee, S. Garg, G. Kaddoum, D. N. K. Jayakody, M. J. Piran and G. Muhammad, "A Trusted Social Network Using Hypothetical Mathematical Model and Decision- Based Scheme," in *IEEE Access*, vol. 9, pp. 4223-4232, 2021, doi: 10.1109/ACCESS.2020.3048077.

[25] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu and V. C. M. Leung, "A Survey on Security Threats and Defensive Techniques of Machine Learning: A Data Driven View," in *IEEE Access*, vol. 6, pp. 12103-12117, 2018, doi: 10.1109/ACCESS.2018.2805680.

[26] S. Jain, V. Sharma and R. Kaushal, "Towards automated real-time detection of misinformation on Twitter," *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Jaipur, India, 2016, pp. 2015-2020, doi: 10.1109/ICACCI.2016.7732347.

[27] E. Osa and O. E. Oghenevbaire, “Comparative Analysis of Machine Learning Models in Computer Network Intrusion Detection”, In *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development*, Lagos, Nigeria, 2022, pp. 1-5, doi: 10.1109/NIGERCON54645.2022.9803175

[28] Z. Chkirbene, R. Camila, A. Erbad, S. Kiranyaz, N. Al-Emadi and M. Hamdi, “Cooperative Machine Learning Techniques for Cloud Intrusion Detection”, In *2021 International Wireless Communications and Mobile Computing*, Harbin City, China pp. 837-842, doi: 10.1109/IWCMC51323.2021.9498809

[29] S. Vanjire and M. Lakshmi, “Behavior-Based Malware Detection System Approach For Mobile Security Using Machine Learning, in *2021 International Conference on Artificial Intelligence and Machine Vision*, Gandhinagar, India, pp. 1-4, doi: 10.1109/AIMV53313.2021.9671009

[30] C. Chen, J Zhong and W. Chen, “Machine Learning-based web security intrusion detection system”, in *2021 3rd International Academic Exchange Conference on Science and Technology Innovation*, Guangzhou, China, pp. 173-177, doi: 10.1109/IAECST54258.2021.9695557

[31] Y. Xiao and J. Qiu, "Exploring and Detecting Opinion Spam on Social Media," *2020 Chinese Automation Congress (CAC)*, Shanghai, China, 2020, pp. 7570-7573, doi: 10.1109/CAC51589.2020.9327093.

[32] H. Afzal and K. Mehmood, "Spam filtering of bi-lingual tweets using machine learning," *2016 18th International Conference on Advanced Communication Technology (ICACT)*, PyeongChang, Korea (South), 2016, pp. 710-714, doi: 10.1109/ICACT.2016.7423530.

[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.

[34] N. A. Ghanem and H. M. Habeeb, "Classifying Suspicious Content on Social Media Networks," *2021 International Conference on Advanced Computer Applications (ACA)*, Maysan, Iraq, 2021, pp. 171-175

[35] K. A. Djaballah, K. Boukhalfa, Z. Ghalem and O. Boukerma, "A new approach for the detection and analysis of phishing in social networks: the case of Twitter," *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*, Paris, France, 2020, pp. 1-8

[36] C. Prashanth, D. Chandrasekaran, B. Pandian, K. Duraipandian, T. Chen and M. Sathiyanarayanan, "Reveal: Online Fake Job Advert Detection Application using Machine Learning," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-6, doi: 10.1109/DELCON54057.2022.9752784.

[37] S. Mitra, T. Tasnim, M. A. R. Islam, N. I. Khan and M. S. Majib, "A Framework to Detect and Prevent Cyberbullying from Social Media by Exploring Machine Learning Algorithms," 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2021, pp. 1-4, doi: 10.1109/IC4ME253898.2021.9768450.

[38] 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.

[39] S. G, D. Chandrasekaran, M. D. Sre and M. Sathiyanarayanan, "Predicting Abnormal User Behaviour Patterns in Social Media Platforms based on Process Mining," 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 2023, pp. 204-209, doi: 10.1109/IITCEE57236.2023.10091025.

Submission Details

**Name of the members that prepared the report:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)**,** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)**,** [Avani Mundra](mailto:amudra@asu.edu)**,** [Justin Young](mailto:jtyoun15@asu.edu)**,** [Anuranjan Dubey](mailto:adubey37@asu.edu)**,** [Rahul Nayak](mailto:rrnayak@asu.edu)**,** [Sangeeth Santhosh](mailto:ssantho9@asu.edu)**,** [Yeshwanth Reddy Chennur](mailto:ychennur@asu.edu)

**Name of the member that submitted this report:** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)  
  
**The report is approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu)**,** [Gautham Vijayaraj](mailto:gvijaya6@asu.edu)