International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 11 Special Issue 02 | 2026

AI-Based Cyberharassment Detection Lab

Prof. Dipti M. Ladhe[1]
[1]*Assistant Professor, Computer Science and Engineering, Padm Dr. V. B. Kolte College of Engineering, Malkapur, Maharashtra,India*

DOI: 10.5281/zenodo.19285072

ABSTRACT

Cyberharassment is a socially significant cybersecurity issue that can cause serious harm to individuals and targeted communities. While prior research has explored cyberharassment detection, AI-based attacks, and societal challenges in detection systems, limited work has focused on experiential learning approaches that engage students with this emerging social cybersecurity problem. This paper presents a hands-on laboratory platform designed to introduce artificial intelligence (AI) concepts and cyberharassment detection to non-computing students with little or no prior AI background. The lab was implemented for social science students at North Carolina A&T University across two academic semesters in 2022. Students completed structured tasks using a detailed lab manual to gain practical experience with AI techniques applied to cyberharassment detection. Learning outcomes were evaluated using pre- and post-surveys measuring students' self-reported AI knowledge and conceptual understanding. The results demonstrate a moderate improvement in students' understanding of AI fundamentals and their application to cyberharassment detection, highlighting the effectiveness of experiential learning for interdisciplinary cybersecurity education.

1.INTRODUCTION

With the rapid growth of social media, cyberharassment—including cyberbullying and cyberhate—has become increasingly prevalent in online interactions. It involves harmful and inappropriate behaviors directed at individuals or specific social groups based on characteristics such as race, gender, sexual orientation, or religion. Cyberharassment is recognized as a critical socially relevant cybersecurity problem due to its severe negative impact on the safety and emotional well-being of victims, particularly teenagers and minority communities. Studies report that a significant percentage of middle and high school students have experienced cyberbullying, with potential consequences including severe psychological distress and increased suicide risk. The emergence of multimodal cyberharassment involving both text and images further complicates detection efforts.

Artificial intelligence (AI) and machine learning (ML) techniques have shown promise in automatically detecting cyberharassment, and major social media platforms have deployed AI-based moderation systems. However, these systems face challenges related to adversarial attacks, robustness, fairness, and ethical concerns, as AI-based detectors may disproportionately affect underrepresented demographic groups and reinforce societal biases.

Given the interdisciplinary nature of cyberharassment, there is a growing need to educate both computing and non-computing students about AI-driven cybersecurity solutions. Experiential learning has been identified as an effective pedagogical approach for teaching complex AI concepts. In this study, a hands-on experiential lab was used over two semesters to teach AI-based cyberharassment detection to non-computing students. Survey results and statistical analysis demonstrate that experiential learning effectively improves students' understanding of AI concepts and their application to socially relevant cybersecurity problems.

2. RELATED WORK

Experiential learning and other active learning approaches are commonly used to teach AI/ML concepts, particularly within engineering and computer science programs. Prior research has highlighted the limitations of traditional data science pedagogy, especially in online learning environments, and shown that experiential learning enhances problem-solving and practical understanding by focusing on real-world challenges rather than isolated methods.

More recent studies have examined large-scale AI education programs designed for learners with diverse academic backgrounds, as well as the use of hands-on and virtual laboratories to improve engagement and self-efficacy in cybersecurity education. Despite these efforts, most existing work focuses on computing students and lacks dedicated experiential AI/ML labs for non-computing learners. This work addresses this gap by developing hands-on experiential labs tailored specifically for non-computing students.

3. DESIGN & DEVELOPMENT OF AI SOCIALLY RELEVANT CYBERHARASSMENT LAB

3.1 Lab Structure

We developed an experiential learning laboratory to introduce non-computing students to the fundamentals of artificial intelligence (AI) and the machine learning (ML) workflow for solving practical problems. The lab is designed to provide hands-on experience while addressing socially relevant cybersecurity issues, particularly cyberharassment.

3.1.1 Learning Goals

The laboratory focuses on three main learning goals:

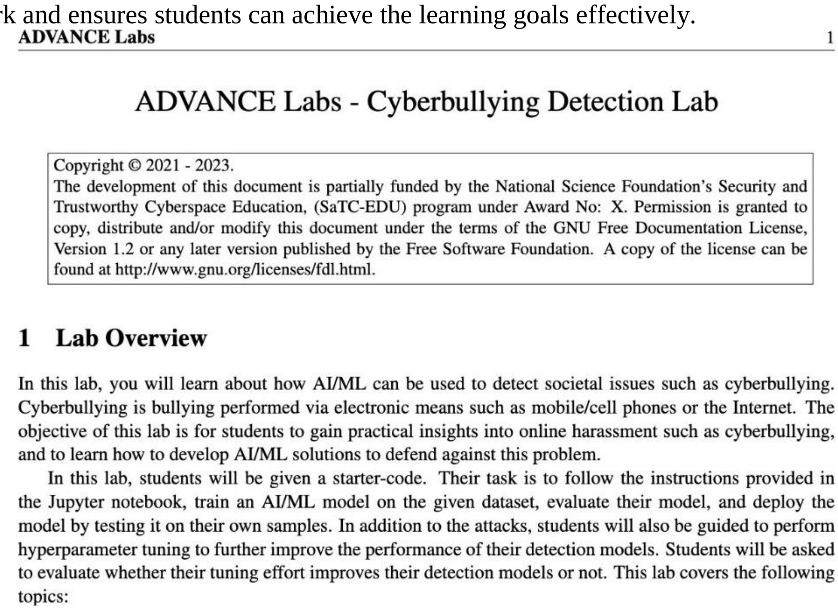
1. Understanding basic AI concepts and the stages of the AI/ML pipeline.
2. Recognizing the nature, consequences, and societal impact of cyberharassment.
3. Gaining practical experience in applying AI to detect and analyze cyberharassment, emphasizing why AI is necessary for large-scale detection where manual methods are insufficient.

These objectives ensure students grasp both the technical and social dimensions of AI-driven cybersecurity solutions.

3.1.2 Lab Structure and Development

The lab was created using an interdisciplinary approach. Initially, AI and cybersecurity specialists designed the technical components, including the cyberharassment detection workflow, considering student backgrounds and intended skills. Social science experts then reviewed the content to enhance clarity, interpretability, and inclusivity, ensuring that students could focus on understanding results rather than complex code implementation. Iterative collaboration between technical and social science teams guided the development of lecture materials, lab assignments, and instructions.

The lab guides students through the entire AI development process, including data preparation, feature engineering, model training, testing, and deployment. Social science insights provide context for understanding cyberharassment and the need for AI intervention. All materials were carefully validated to make them approachable for learners without prior computing experience. A detailed instruction manual supports independent work and ensures students can achieve the learning goals effectively.


ADVANCE Labs 1

ADVANCE Labs - Cyberbullying Detection Lab

Copyright © 2021 - 2023.
The development of this document is partially funded by the National Science Foundation's Security and Trustworthy Cyberspace Education, (SaTC-EDU) program under Award No: X. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation. A copy of the license can be found at http://www.gnu.org/licenses/fdl.html.

1 Lab Overview

In this lab, you will learn about how AI/ML can be used to detect societal issues such as cyberbullying. Cyberbullying is bullying performed via electronic means such as mobile/cell phones or the Internet. The objective of this lab is for students to gain practical insights into online harassment such as cyberbullying, and to learn how to develop AI/ML solutions to defend against this problem.

In this lab, students will be given a starter-code. Their task is to follow the instructions provided in the Jupyter notebook, train an AI/ML model on the given dataset, evaluate their model, and deploy the model by testing it on their own samples. In addition to the attacks, students will also be guided to perform hyperparameter tuning to further improve the performance of their detection models. Students will be asked to evaluate whether their tuning effort improves their detection models or not. This lab covers the following topics:

Fig. 1. A screenshot illustrating Lab 1 instruction manual

3.1.3 Lab Delivery

The laboratory is delivered using a three-stage approach, combining instructor-led lectures, guided hands-on exercises, and independent student work.

In the first stage, students attend a lecture prepared by the research team to provide foundational knowledge. This session introduces the concepts of cyberharassment, AI fundamentals, the AI development workflow, and the rationale for using AI in cyberharassment detection.

The second stage consists of guided hands-on exercises hosted on the Google Colab platform (Figure 2). This interactive environment allows students to directly implement and test AI models for cyberharassment detection.

Colab's combination of text and code cells enables step-by-step instructions alongside executable Python code, while visual aids (e.g., Figure 3) support conceptual understanding.

The final stage emphasizes independent practice, allowing students to consolidate their learning, experiment with AI models, and reinforce the practical application of theoretical concepts. This structured delivery ensures that students not only understand AI techniques but also appreciate their relevance in addressing socially impactful cybersecurity challenges.

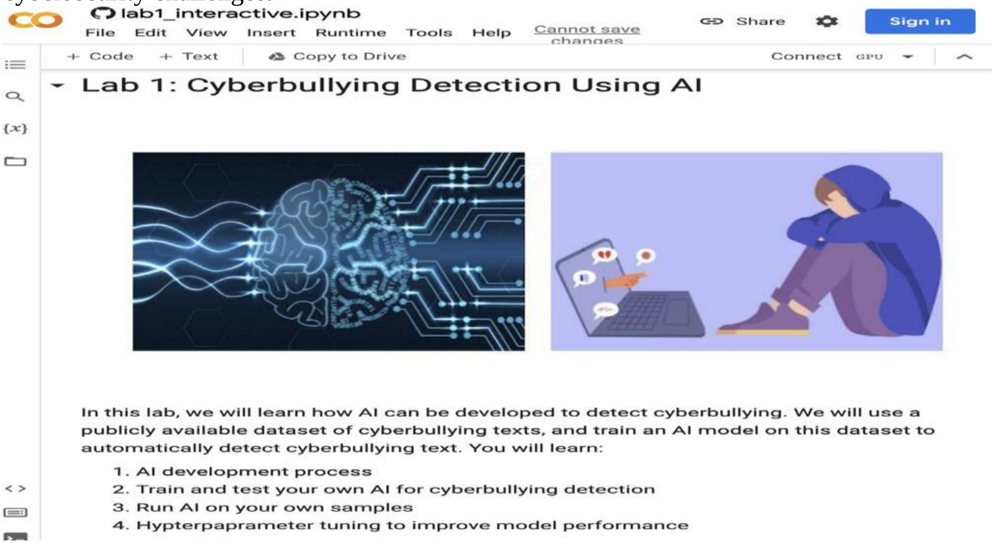


Fig. 2. A screenshot illustrating the lab interface

The lab workflow on Google Colab requires sequential execution of code cells, and students are instructed to follow this order to ensure proper functioning of the AI model. The lab manual explicitly guides students through these dependencies to prevent execution errors and to reinforce structured problem-solving.

4.METHODS

4.1. Survey Design

To assess the effectiveness of the laboratory, we designed pre- and post-lab surveys to collect both quantitative and qualitative feedback. Students evaluated their knowledge of AI and machine learning techniques for cyberharassment detection using a five-point Likert scale, where 1 indicated "Proficient/Strongly Agree" and 5 indicated "None/Strongly Disagree." The post-survey additionally captured students' perceptions of the lab experience from multiple perspectives and included three open-ended questions to identify which aspects of the lab were most helpful, the challenges encountered, and suggestions for improving the learning experience.

4.2. Data Collection

Prior to data collection, approval was obtained from the institution's Institutional Review Board (IRB). Participation was voluntary, and students were assured that all responses would remain confidential and accessible only to the research team. The pre-survey was administered before the lab session to gauge baseline knowledge, and the post-survey was administered afterward to measure learning outcomes. Surveys were conducted using the Qualtrics platform [37].

4.3. Data Analysis

Quantitative data were analyzed by comparing mean scores between pre- and post-surveys. Statistical significance was determined using a sample t-test at a 0.05 significance level. A p-value below 0.05 was interpreted as indicating a significant difference. All analyses were performed using SciPy, an open-source Python library for scientific computing.

4.4. AI Model for Cyberharassment Detection

The laboratory incorporates a pre-trained BERT transformer model , accessed via HuggingFace, for cyberharassment detection. Students interact with the model by executing a designated code cell in Google Colab, after which the model becomes available throughout the platform for experimentation and testing.

5.CONCLUSION

We developed an AI-based cybersecurity lab for non-computing students, providing hands-on experience in detecting cyberharassment. The Fall 2022 iteration, improved based on student feedback, significantly enhanced understanding of AI and machine learning concepts. Clear instructions, real-world problem contexts, and interdisciplinary collaboration were key to engagement and learning. Future work will focus on refining the lab and developing additional modules on multimodal detection, adversarial attacks, and bias mitigation, extending applicability to computing students.

6. REFERENCES

[1] Y. Velaj, D. Dolezal, R. Ambros, C. Plant, and R. Motschnig, "Designing a data science course for non-computer science students: Practical considerations and findings," pp. 1–9, 2022.
[2] A. Y. Kolb and D. A. Kolb, "Learning styles and learning spaces: Enhancing experiential learning in higher education," *Academy of Management Learning & Education*, vol. 4, no. 2, pp. 193–212, 2005.
[3] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*. FT Press, 2014.
[4] A. Barman, S. Chen, A. Chang, and G. Allen, "Experiential learning in data science through a novel client-facing consulting course," pp. 1–9, 2022.
[5] G. I. Allen, "Experiential learning in data science: Developing an interdisciplinary, client-sponsored capstone program," pp. 516–522, 2021.
[6] A. F. Salazar-Gomez, A. Bagiati, N. Minicucci, K. D. Kennedy, X. Du, and C. Breazeal, "Designing and implementing an AI education program for learners with diverse background at scale," pp. 1–8, 2022.
[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.