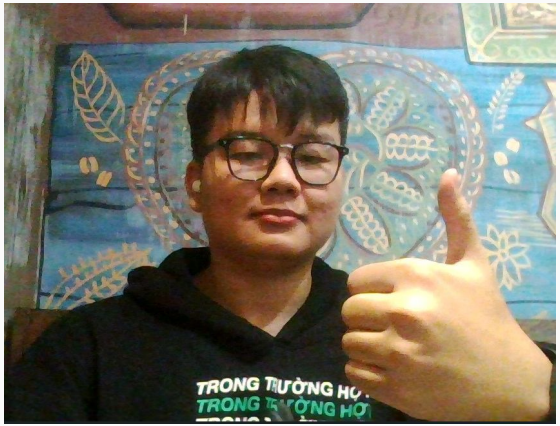
NGHIÊN CỨU PHƯƠNG PHÁP XÂY DỰNG HỆ THỐNG PHÁT HIỆN VÀ NGĂN CHẶN TẤN CÔNG SỬ DỤNG TRÍ TUỆ NHÂN TẠO ĐỂ CHỐNG LẠI CÁC CUỘC TẤN CÔNG ĐỐI KHÁNG

Lê Chí Đại - 240202019

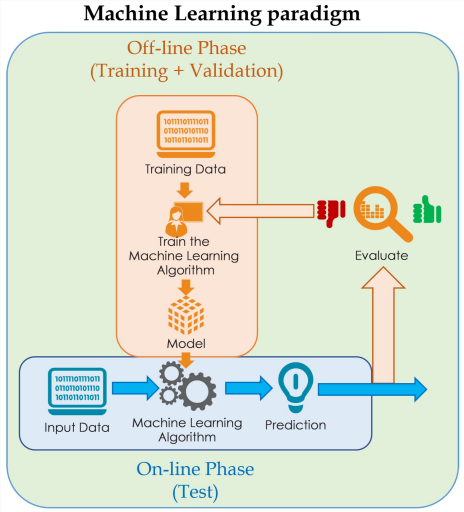
Tóm tắt

- Lớp: CS2205.FEB2025
- Link Github của nhóm: https://github.com/akabeomip/CS2205_IDPS_ADVERSARIAL
- Link YouTube video: https://www.youtube.com/watch?v=3ngyIVMAKro
- Lê Chí Đại - 240202019



Giới thiệu

Đề tài "Xây dựng hệ thống AI-base IDPS để phát hiện các cuộc tấn công đối kháng" tập trung vào việc ứng dụng trí tuệ nhân tạo (AI) nhằm nâng cao khả năng phát hiện và phản ứng của IDPS. Bằng cách sử dụng các thuật toán học máy và học sâu, hệ thống này có khả năng phân tích dữ liệu lớn và nhận diện các mẫu hành vi bất thường, từ đó phát hiện những cuộc tấn công mà các phương pháp truyền thống không thể nhận diện.

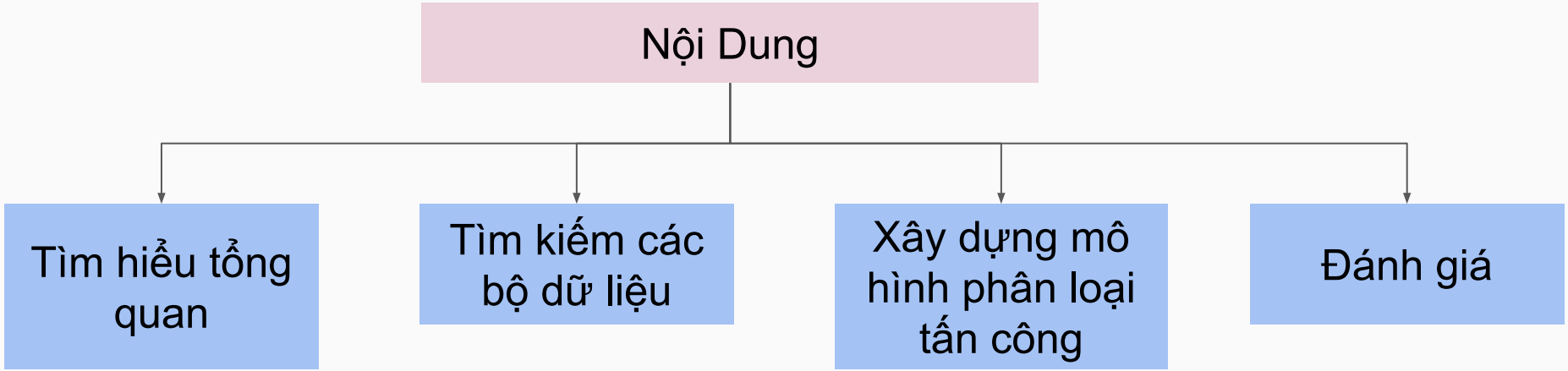


Hình 1. Minh họa hệ thống bài toán

Mục tiêu

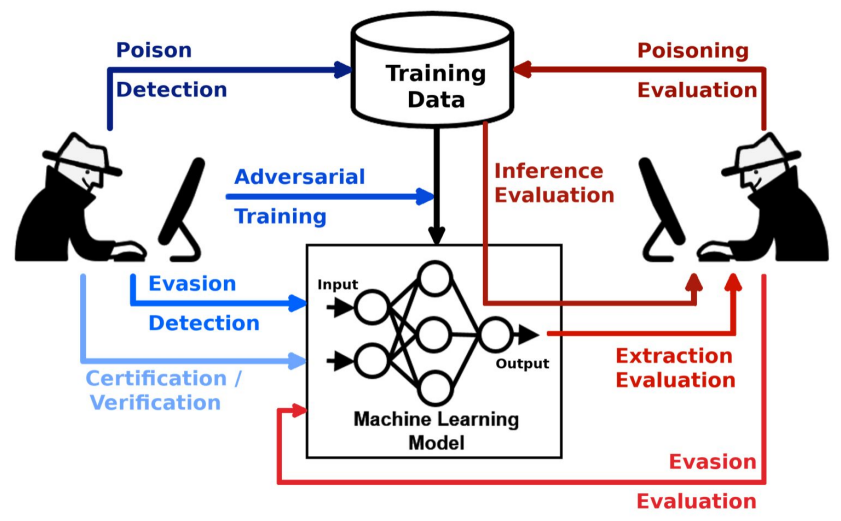
- Xây dựng tập dữ liệu dành riêng cho bài toán phát hiện xâm nhập bình thường và phát hiện xâm nhập đối kháng.
- Xây dựng mô hình phát hiện xâm nhập.
- Kết quả đánh giá và so sánh của mô hình vừa xây dựng được trên hai tập dữ liệu tấn công cổ điển và dữ liệu tấn công đối kháng.

Nội dung và Phương pháp



Nội dung và Phương pháp

Các cuộc tấn công đối kháng có hiệu suất rất cao khi tấn công vào các hệ thống phát hiện và ngăn chặn tấn công sử dụng trí tuệ nhân tạo [2]



Hình 2. Minh họa đầu vào đầu ra của module tiền xử lý

Kết quả dự kiến

- Xây dựng được mô hình phân loại các cuộc tấn công cổ điển.
- Xây dựng được mô hình phân loại phát hiện được các mẫu tấn công đối kháng
- Kết quả của mô hình đề xuất có kết quả tốt trên các bộ dữ liệu cổ điển và phát hiện được các mẫu đối kháng.

Tài liệu tham khảo

[1]. Akhil Krishna, Ashik Lal M.A., Athul Joe Mathewkutty, Dhanya Sarah Jacob, M. Hari (2020). Intrusion Detection and Prevention System Using Deep Learning

[2]. Igino Corona, Giorgio Giacinto, Fabio Roli (2013). Adversarial attacks against intrusion detection systems: Taxonomy, solutions and open issues

[3]. Afnan Alotaibi, Murad A. Rassam (2023). Adversarial Machine Learning Attacks against Intrusion Detection Systems: A Survey on Strategies and Defense

[4]. Chaoyun Zhang, Xavier Costa-Pérez, Paul Patras (2022). Adversarial Attacks Against Deep Learning-Based Network Intrusion Detection Systems and Defense Mechanisms