# NÂNG CAO HIỆU QUẢ PHÁT HIỆN MÃ ĐỘC BẰNG GIẢI PHÁP HỌC SÂU DỰA TRÊN DỮ LIỆU HÀNH VI TỪ CAPE SANDBOX

Dương Quốc Cường - 240202003

# Tóm tắt

- Lớp: CS2205.CH183
- Link Github của nhóm: https://github.com/cyrus131/CS2205.CH183/blob/main/EnhancingMalwareDetectionEfficiencyWithDeepLearningBasedOnBehavioralDataFromTheCapeSandbox.pdf
- Link YouTube video: https://youtu.be/63svt-cnFqk
- Ảnh + Họ và Tên của các thành viên:
  - Dương Quốc Cường

# Giới thiệu

- Phần mềm độc hại ngày càng tinh vi, làm giảm hiệu quả của các phương pháp truyền thống.
- Nghiên cứu đề xuất sử dụng BERT để phân loại mã độc dựa trên hành vi từ CAPE Sandbox. Nhờ khả năng học ngữ cảnh hai chiều, BERT nắm bắt mối quan hệ giữa API call và hành vi mã độc.
- Mục tiêu đề ra là so sánh hiệu quả của BERT với Random Forest và SVM, đề xuất giải pháp nâng cao phát hiện mã độc.

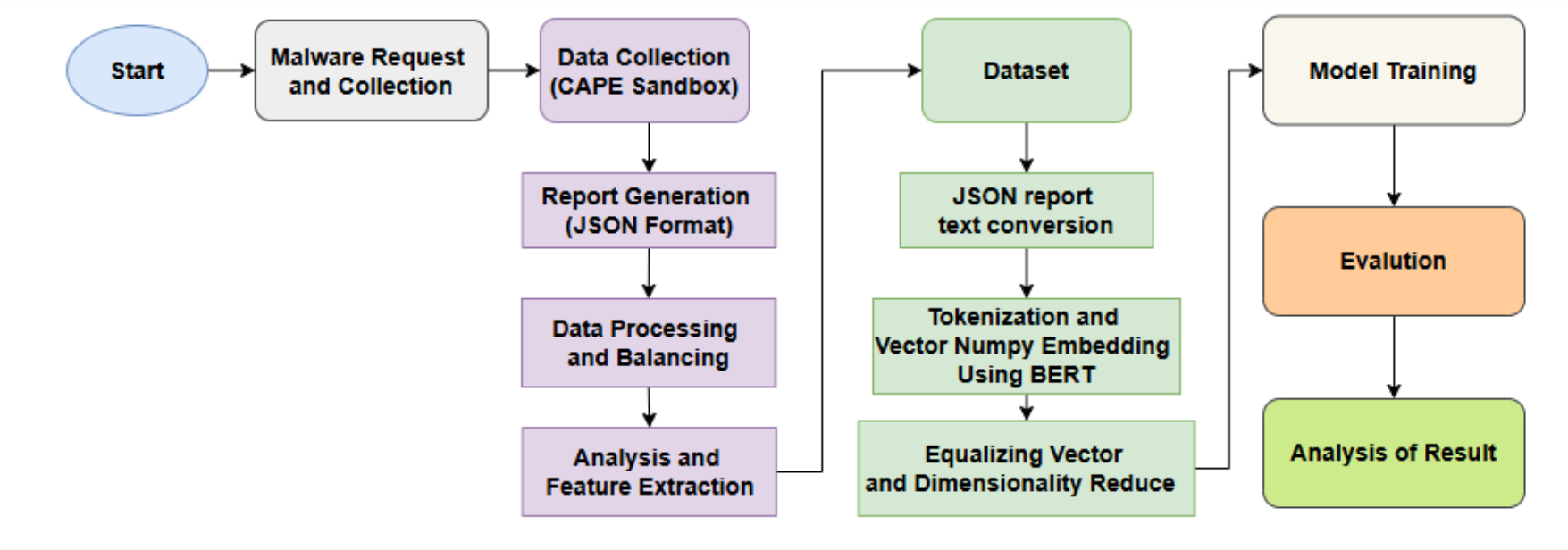
# Mục tiêu

- Xây dựng bộ dữ liệu hành vi mã độc từ CAPEv2 Sandbox, bao gồm API calls, tiến trình, registry, tập tin, và kết nối mạng.

- Huấn luyện mô hình BERT và so sánh với các thuật toán truyền thống (Random Forest, SVM) để đánh giá hiệu suất phát hiện mã độc.

- Tối ưu hóa mô hình, đề xuất ứng dụng mô hình trong hệ thống giám sát an ninh mạng nhằm cải thiện khả năng phòng thủ trước mã độc chưa từng biết.

# Nội dung và Phương pháp

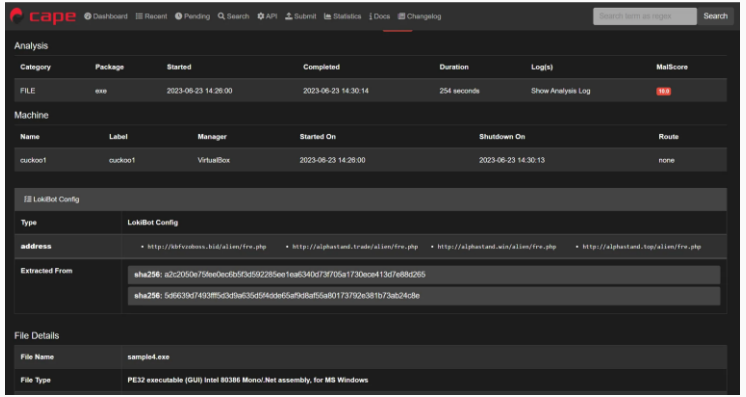


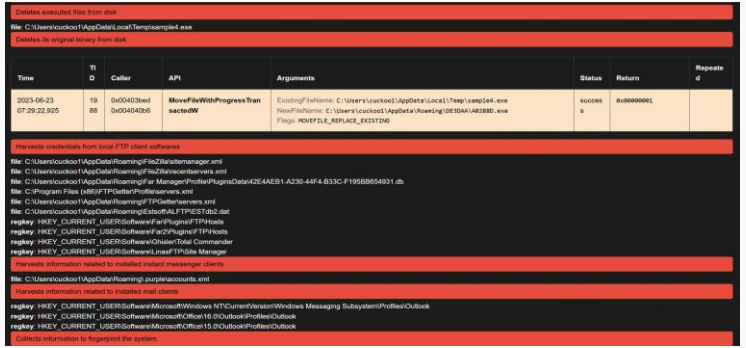

*Hình. Quy trình tổng thể*

# Nội dung và Phương pháp

## Giai đoạn 1. Thu thập và tiền xử lý dữ liệu

- Thu thập mã độc từ VirusTotal, các nguồn lưu trữ mã độc.
- Chạy các mẫu mã độc trên CAPEv2.
- Tiền xử lý dữ liệu, chuẩn hóa và tạo chuỗi token.

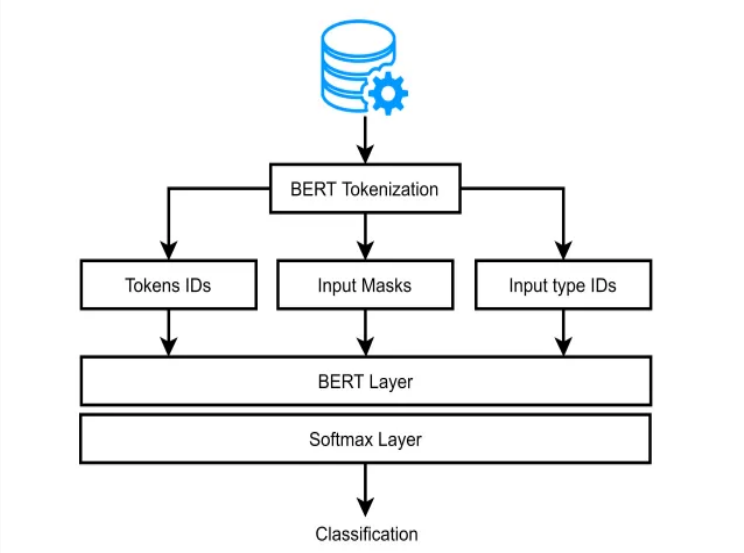


# Nội dung và Phương pháp

## Giai đoạn 2. Xây dựng mô hình BERT

Dữ liệu hành vi sau khi xử lý được chuyển đổi thành dạng chuỗi văn bản, sau đó được đưa vào mô hình BERT đã được tinh chỉnh




- Mô hình BERT làm phần rút trích đặc trưng từ chuỗi hành vi.
- Lớp phân loại đầu ra với softmax hoặc sigmoid để dự đoán mã độc.

# Nội dung và Phương pháp

**Giai đoạn 3. Huấn luyện và tối ưu**

- Sử dụng AdamW, dropout, weight decay.
- Tinh chỉnh siêu tham số như learning rate, batch size.

**Giai đoạn 4. Đánh giá và so sánh**

- Đánh giá bằng Accuracy, Precision, Recall, F1-score.
- So sánh với SVM, Random Forest, LSTM.

# Kết quả dự kiến

- Xây dựng mô hình BERT tối ưu hóa đầu vào gồm tiến trình, tập tin, registry, kết nối mạng và hành động hệ thống.
- Cải thiện độ chính xác so với Random Forest, SVM, đồng thời nâng cao khả năng phát hiện mã độc zero-day.
- Triển khai thử nghiệm, tích hợp vào hệ thống giám sát an ninh mạng, hỗ trợ phát hiện và phân tích mã độc theo thời gian thực.

# Tài liệu tham khảo

[1] Devlin, J., Chang, M., Lee, K., & Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv (2018).
[2] Ki, Y., Kim, E., & Kim, H. K. A novel approach to detect malware based on API call sequence analysis. Int. J. Distributed Sensor Networks, 11 (2015), 659101.
[3] Kolosnjaji, B., Zarras, A., Webster, G., & Eckert, C. Deep learning for classification of malware system call sequences. Australasian Joint Conference on Artificial Intelligence (2016), pp. 137-149.
[4] Demirkıran, F., Çayır, A., Ünal, U., & Dağ, H. An ensemble of pre-trained transformer models for imbalanced multiclass malware classification. arXiv preprint arXiv:2112.13236 (2021).
[5] Alqahtani, S., & Jones, J. A. Dynamic malware analysis using machine learning techniques. IEEE (2021), pp. 5199–5202.