THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
  https://www.youtube.com/watch?v=d0Om41uPnN4
- Link slides (dạng .pdf đặt trên Github của nhóm):
  https://github.com/locnthIA/CS2205.NOV2024/blob/main/Slide_Nghien%20cuu%20giai%20phap%20phat%20hien%20som%20lua%20dao%20giong%20noi%20su%20dung%20Progressive%20learning.pdf
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*
- *Lớp Cao học, mỗi nhóm một thành viên*

- Họ và Tên: Nguyễn Thanh Hoài Lộc
- MSSV: 240202009

- Lớp: CS2205.NOV2024
- Tự đánh giá (điểm tổng kết môn): 9/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 6
- Link Github:
  https://github.com/locnthIA/CS2205.NOV2024

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

NGHIÊN CỨU GIẢI PHÁP PHÁT HIỆN SỚM LỪA ĐẢO GIỌNG NÓI SỬ DỤNG PROGRESSIVE LEARNING

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

EARLY DETECTION OF VOICE SCAM USING PROGRESSIVE LEARNING APPROACH

TÓM TẮT *(Tối đa 400 từ)*

Sự gia tăng đáng báo động của các cuộc gọi lừa đảo đang trở thành thách thức lớn đối với an ninh mạng tại Việt Nam. Phương pháp phát hiện truyền thống thường phản ứng chậm và thiếu chính xác do đặc thù phức tạp của hình thức tấn công này. Dựa trên tiềm năng của kỹ thuật Progressive Learning trong việc thích nghi và cải thiện liên tục, nghiên cứu này đề xuất một giải pháp mới cho phép phát hiện sớm các cuộc gọi lừa đảo. Chúng tôi xây dựng hai nguồn dữ liệu độc đáo: bộ ghi âm cuộc gọi được chú thích chuyên sâu và tập dữ liệu tổng hợp từ hệ thống giám sát. Kết quả thực nghiệm cho thấy hệ thống có khả năng cảnh báo chính xác trong 30 giây đầu cuộc gọi, mở ra hướng tiếp cận mới trong bảo vệ người dùng.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Lừa đảo qua cuộc gọi điện thoại đang là một trong những hình thức tấn công phổ biến nhất tại Việt Nam. Với sự phát triển của công nghệ, các thủ đoạn lừa đảo ngày càng tinh vi và khó phát hiện hơn. Các phương pháp truyền thống dựa trên phân tích hậu kiểm thường không đủ nhanh để ngăn chặn thiệt hại, trong khi các giải pháp phát hiện thời gian thực lại gặp khó khăn với tỷ lệ cảnh báo sai cao.

Progressive Learning, một kỹ thuật học máy tiên tiến, cho phép mô hình liên tục học

và thích nghi từ dữ liệu mới mà vẫn duy trì được kiến thức đã học trước đó. Trong lĩnh vực phát hiện bất thường, Progressive Learning đã chứng minh hiệu quả vượt trội trong việc cải thiện độ chính xác và giảm thời gian phát hiện. Đặc biệt, khả năng học tiến triển theo thời gian thực của phương pháp này mở ra tiềm năng lớn trong việc phát hiện sớm các cuộc gọi lừa đảo.

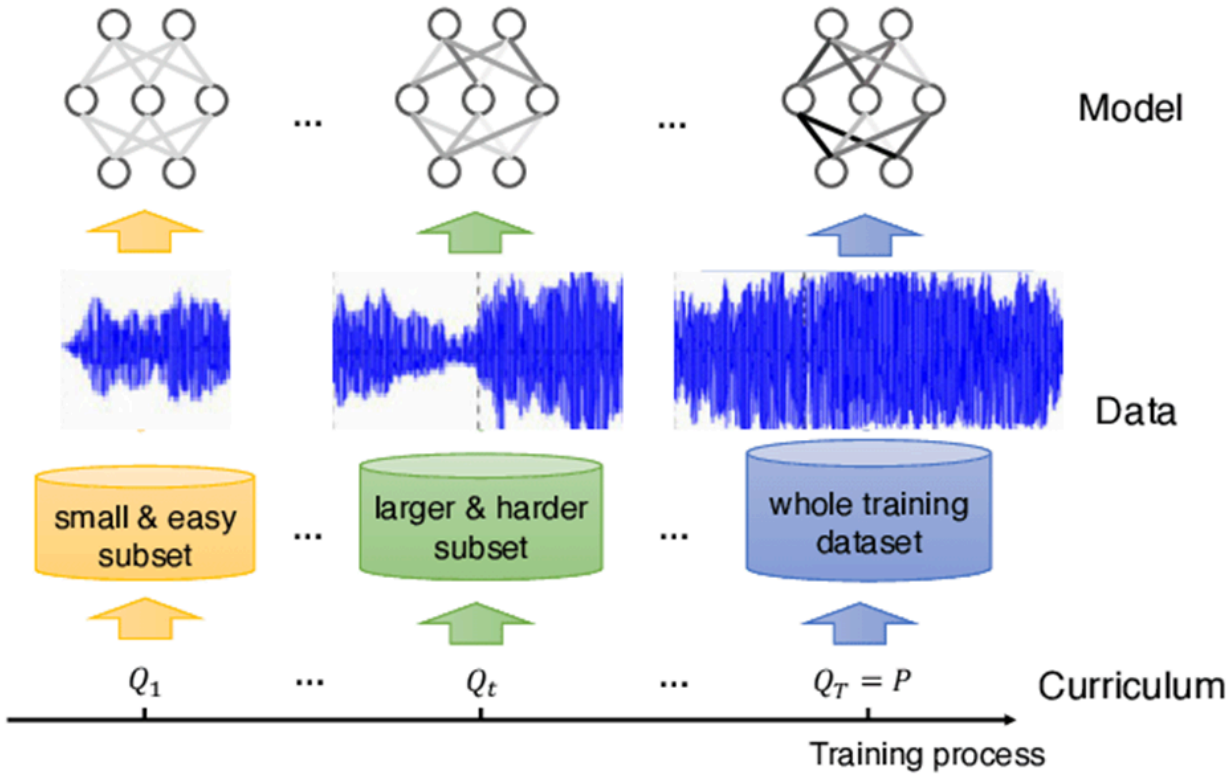
Progressive Learning trong phát hiện lừa đảo giọng nói mở ra một hướng tiếp cận mới, cho phép hệ thống nhanh chóng nhận diện các dấu hiệu đáng ngờ ngay từ những giây đầu tiên của cuộc gọi. Phương pháp này đặc biệt hữu ích trong việc giảm thiểu thiệt hại cho người dùng, đồng thời cung cấp khả năng thích nghi với các hình thức lừa đảo mới.

MỤC TIÊU *(Viết trong vòng 3 mục tiêu)*

1. Xây dựng hệ thống thông minh tích hợp Progressive Learning kết hợp Deep Learning, cho phép cảnh báo nhanh chóng các cuộc gọi đáng ngờ trong vòng 30 giây đầu tiên.
2. Thiết lập kho dữ liệu chuyên biệt về cuộc gọi lừa đảo với chú thích chi tiết, phục vụ nghiên cứu và phát triển các giải pháp bảo vệ người dùng.
3. Nâng cao hiệu quả phát hiện thông qua cơ chế tự học liên tục, cho phép hệ thống thích nghi với các chiến thuật lừa đảo mới.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung: Nghiên cứu tập trung vào việc phát triển mô hình Progressive Learning cho phát hiện sớm lừa đảo giọng nói. Mô hình được đề xuất kết hợp các kỹ thuật xử lý âm thanh thời gian thực với deep learning framework cho phép học và cập nhật liên tục. Khác với các nghiên cứu trước chỉ sử dụng mô hình tĩnh, chúng tôi khai thác khả năng học tiến triển để cải thiện hiệu suất theo thời gian.



Gồm ba giai đoạn chính, kết hợp xử lý âm thanh và Progressive Learning để phát hiện sớm lừa đảo giọng nói:

- Giai đoạn 1: Xử lý tín hiệu âm thanh
    - Mục đích: Chuẩn hóa và trích xuất đặc trưng từ dữ liệu cuộc gọi thô.
    - Cách thực hiện:
        - Xử lý tín hiệu số để loại bỏ nhiễu và chuẩn hóa âm thanh đầu vào.
        - Phân đoạn cuộc gọi theo khoảng thời gian để tập trung vào 30 giây đầu tiên.
        - Trích xuất đặc trưng âm thanh bao gồm đặc trưng ngôn ngữ và phi ngôn ngữ.
        - Đầu ra: Véc-tơ đặc trưng đã được chuẩn hóa, sẵn sàng cho phân tích bằng Deep Learning.
- Giai đoạn 2: Mô hình Deep Learning
    - Mục đích: Phân tích các đặc trưng âm thanh để phát hiện dấu hiệu lừa đảo.

- Cách thực hiện:
    - Kiến trúc module: Chia hệ thống thành các thành phần độc lập chuyên biệt.
    - Sử dụng các mạng neural sâu để phát hiện mẫu bất thường trong giọng nói.
    - Tích hợp các thành phần phân tích ngôn ngữ và ngữ điệu để nhận diện hành vi lừa đảo.
    - Đầu ra: Xác suất cuộc gọi là lừa đảo dựa trên các mẫu đã học.
- Giai đoạn 3: Progressive Learning Framework
    - Mục đích: Cho phép hệ thống liên tục học và thích nghi với các chiến thuật lừa đảo mới.
    - Cách thực hiện:
        - Xây dựng framework học tăng tiến để duy trì kiến thức cũ trong khi tích hợp kiến thức mới.
        - Học tăng cường để liên tục cập nhật mô hình dựa trên phản hồi từ hệ thống giám sát.
        - Cơ chế thích nghi cho phép điều chỉnh ngưỡng cảnh báo dựa trên dữ liệu mới.
        - Nhiệm vụ phụ trợ: Phân loại các loại lừa đảo để cung cấp thông tin chi tiết hơn về mối đe dọa.

KẾT QUẢ MONG ĐỢI

- Hệ thống thông minh có khả năng phát hiện cuộc gọi lừa đảo trong 30 giây với độ tin cậy cao.
- Kho dữ liệu chuyên sâu với trên 10,000 mẫu cuộc gọi được phân tích và chú thích.
- Giải pháp hoàn chỉnh sẵn sàng triển khai với giao diện API và hệ thống giám sát.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

[1] M. K. Bae, S. Kim, và H. Ko, "Class-Incremental Learning for Sound Event Localization and Detection," *arXiv preprint arXiv:2411.12830*, 2024.

[2] N. H. Tuấn, "Cơ sở toán và MFCCs – Trích xuất đặc trưng âm thanh," *Tạp chí Khoa học Trường Đại học Sư phạm TP Hồ Chí Minh*, tập 20, số 4, trang 55–67, 2024.

[3] J. Zhang, A. Kumar, và S. Tan, "Progressive Continual Learning for Spoken Keyword Spotting," trong *Kỷ yếu Hội nghị ICASSP 2024*, trang 3255–3259.

[4] L. Wang và K. Patel, "Towards Robust Few-shot Class Incremental Learning in Audio Classification using Contrastive Representation," *arXiv preprint arXiv:2407.19265*, 2024.

[5] H. Lee, J. Park, và M. Chen, "Audio-Visual Class-Incremental Learning," *arXiv preprint arXiv:2308.11073*, 2023.