Bộ môn An toàn Thông tin – Khoa Mạng máy tính và Truyền thông
Trường Đại học Công nghệ Thông tin – Đại học Quốc gia thành phố Hồ Chí Minh

6

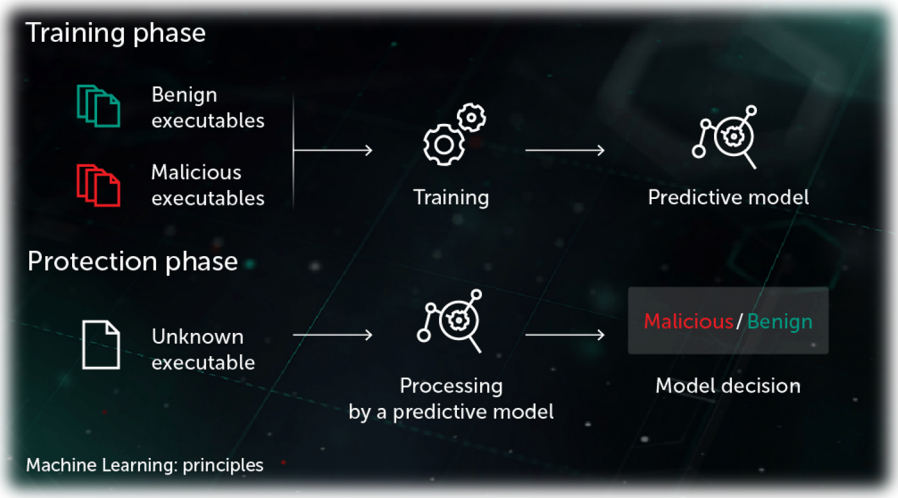
Lab

PHỤC VỤ MỤC ĐÍCH GIÁO DỤC
FOR EDUCATIONAL PURPOSE ONLY

Machine Learning based Intrusion Detection System

Python

Thực hành môn Phương pháp học máy trong an toàn thông tin



Tháng 6/2025

Lưu hành nội bộ

<Nghiêm cấm đăng tải trên internet dưới mọi hình thức>

Mọi góp ý về tài liệu, vui lòng gửi về email inseclab@uit.edu.vn

A. TỔNG QUAN

1. Thời gian thực hành

- Thực hành tại lớp: **1 giờ**.

2. Môi trường thực hành

Google Colab (https://research.google.com/colaboratory/)

B. THỰC HÀNH

Michael Learning là một thanh niên trẻ đam mê lĩnh vực bảo mật và an ninh mạng. Anh ta là một lập trình viên giỏi và luôn tìm kiếm thách thức mới để áp dụng kiến thức của mình. Một ngày nọ, **Michael** nhận được một cơ hội thú vị từ một công ty an ninh mạng hàng đầu.

Công ty đang đối mặt với vấn đề ngày càng phức tạp của các cuộc tấn công mạng và muốn phát triển một mô hình phát hiện xâm nhập mới để bảo vệ hệ thống của họ. Họ đã nghe về tiềm năng và tài năng của **Michael** và quyết định giao cho anh ta nhiệm vụ quan trọng này.

Cuộc hành trình của **Michael** bắt đầu. Anh ta bắt đầu nghiên cứu sâu về các phương pháp phát hiện xâm nhập và các kỹ thuật phân tích bất thường. Anh ta khám phá ra rằng việc xây dựng một mô hình phát hiện xâm nhập hiệu quả đòi hỏi một sự kết hợp hoàn hảo giữa việc hiểu rõ các mẫu tấn công phổ biến và khả năng phát hiện những hành vi không bình thường.

Michael tiếp tục đặt ra các câu hỏi quan trọng để định hình mô hình của mình: Làm thế nào để xác định các dấu hiệu của một cuộc tấn công? Làm thế nào để phân biệt giữa một hoạt động đáng ngờ và một hành vi thông thường? Làm thế nào để phát hiện những biểu hiện tiềm năng của một cuộc tấn công chưa từng xảy ra trước đó?

Qua nỗ lực không ngừng, **Michael** đã xây dựng một mô hình phát hiện xâm nhập. Mô hình của anh ta sử dụng các thuật toán máy học tiên tiến để phân tích dữ liệu mạng và phát hiện các hoạt động không bình thường. Điều này giúp công ty giảm thiểu rủi ro xâm nhập mạng và đảm bảo an toàn cho hệ thống của họ.

Cuối cùng, công ty vô cùng hài lòng với kết quả của **Michael**. Anh ta đã chứng minh khả năng của mình trong việc xây dựng mô hình phát hiện xâm nhập đột phá và đóng góp quan trọng cho an ninh mạng của công ty.

Câu chuyện về **Michael** trở thành một nguồn cảm hứng cho những người trẻ khác trong lĩnh vực an ninh mạng. Anh ta đã chứng minh rằng với kiến thức, đam mê và sự kiên nhẫn, mọi thách thức đều có thể vượt qua.

Các bạn hãy như **Michael xây dựng một mô hình phát hiện xâm nhập mạng** (Binary Classification) với tập dữ liệu huấn luyện cung cấp sẵn (**X_Train.csv** và **Y_Train.csv**) và đánh giá mô hình đó với tập dữ liệu kiểm thử (**X_test.csv** và **Y_test.csv**).

Cùng với đó **Michael** cũng mang đến cho các bạn một bài thơ, hãy nắm bắt những **từ** bạn cho là **chìa khoá** để tìm ra bí ẩn có trong tập **Key_feature.csv**:

"Trong vũ trụ triền miên, đong đầy bí mật,
An toàn thông tin, là mối quan tâm lớn.
Số 200, khung cửa rộng mở trước mắt,
50 bước chạm tới sự an lành bình yên.

Không là đen, **một là trắng**, tuyệt vời sáng ngời,
Bảo mật dữ liệu, không để bị xâm phạm.
Với công nghệ tiên tiến, mãi mãi đáng tin cậy,
Giữ kín bí mật, khỏi rơi vào tay tham lam.

Nguyên tắc đạo đức, là chìa khóa vàng,
Trung thực, trách nhiệm, đường đi không lạc lối.
Số 200, biểu tượng an toàn và đáng tin,
50, đường dẫn tới thế giới không gian vô tận.

Công nghệ bảo mật, là gương phản chiếu,
Phòng chống tấn công, truy cập trái phép.
Bảo vệ thông tin, từ khách hàng đến doanh nghiệp,
Tạo nên một môi trường an toàn và tuyệt vời.

200 và 50, số đẹp trong an ninh,
Dẫn đường cho sự phát triển vững bền.
Không là đen, một là trắng, lòng tin trọn vẹn,
An toàn thông tin, hạnh phúc đến tận cùng trời xanh."

HẾT

Chúc các bạn hoàn thành tốt!