THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút): https://youtu.be/zs60jXXCqDY
- Link slides (dạng .pdf đặt trên Github của nhóm): CS2205.CH183/Huyền Nguyễn Ngọc - CS2205.NOV2024.DeCuong.FinalReport.Template.Slide.pdf at main · HuyenNgocNguyen18102001/CS2205.CH183
- *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 Ngọc Huyền
- MSSV: 230202007

- Lớp: **CS2205.CH183**
- Tự đánh giá (điểm tổng kết môn): 9.0/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 5
- Link Github:

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

ỨNG DỤNG MACHINE LEARNING TRONG HỆ THỐNG IT SERVICE MANAGEMENT (ITSM): TỰ ĐỘNG PHÁT HIỆN VÀ XỬ LÝ SỰ CỐ BẰNG SERVICENOW VÀ SPLUNK

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

APPLYING MACHINE LEARNING IN IT SERVICE MANAGEMENT (ITSM): AUTOMATED INCIDENT DETECTION AND RESOLUTION USING SERVICENOW AND SPLUNK

TÓM TẮT

Việc quản lý và giám sát hệ thống là một thách thức quan trọng đối với các tổ chức và doanh nghiệp. Hệ thống thông tin phải đảm bảo tính ổn định và hiệu suất cao để giảm thiểu thời gian gián đoạn dịch vụ. Bài nghiên cứu này tập trung vào việc xây dựng một hệ thống tự động xử lý sự cố (incident) sử dụng ServiceNow [1] để quản lý quy trình ITSM (IT Service Management) và Splunk [2] để thu thập, phân tích dữ liệu sự cố.

Hệ thống đề xuất sẽ giúp tự động phát hiện sự cố, gửi cảnh báo và xử lý vấn đề một cách nhanh chóng, giảm thiểu sự can thiệp thủ công bằng việc ứng dụng các thuật toán học máy. Nghiên cứu cũng phân tích các giải pháp hiện tại như Accenture MyWizard [3], Freshdesk [4] và các phương pháp khắc phục nhược điểm nhằm cải thiện tính hiệu quả của hệ thống.

GIỚI THIỆU

Giám sát và quản lý hệ thống công nghệ thông tin (CNTT) đóng vai trò quan trọng trong việc đảm bảo tính ổn định và liên tục của các dịch vụ. Khi xảy ra sự cố, việc kiểm tra và khắc phục thủ công không chỉ tốn thời gian mà còn dễ dẫn đến sai sót, làm gián đoạn hoạt động kinh doanh.

Hiện nay, Machine Learning (ML) đang được ứng dụng rộng rãi trong quản lý dịch vụ CNTT (IT Service Management - ITSM) để tự động hóa quá trình phân loại, chẩn đoán và xử lý sự cố. Theo nghiên cứu của Dmitry Zuev và cộng sự [5], các mô hình ML giúp cải thiện độ chính xác trong việc phân loại ticket và giảm thời gian xử lý sự cố. Tác giả chỉ ra rằng việc áp dụng học máy giúp giảm đáng kể thời gian phản hồi và tối ưu hóa quy trình xử lý ticket. Một số thuật toán phổ biến như **Naive Bayes, Random Forest, và Gradient Boosting Decision Trees** đã được sử dụng để phân loại ticket và dự đoán mức độ ưu tiên trong hệ thống ITSM với độ chính xác lên đến 82%.

Trong khi đó, nghiên cứu của Rajeev Gupta và cộng sự [6] cho thấy **Xử lý ngôn ngữ tự nhiên (NLP)** có thể giúp phân tích và chuẩn hóa dữ liệu phi cấu trúc từ các ticket sự cố. Sau đó, các mô hình ML như **Naive Bayes, SVM, K-Means và TF-IDF** được sử dụng để phân loại sự cố, phát hiện các sự cố lặp lại và dự đoán nguyên nhân gốc rễ (root cause). Các mô hình **Decision Trees** có thể dự đoán các bước xử lý dựa trên dữ liệu sự cố trước đó, trong khi **học tăng cường**

(Reinforcement Learning) tối ưu hóa quy trình xử lý bằng cách học từ phản hồi của các kỹ sư hệ thống. Nghiên cứu này đạt độ chính xác 82% trong dự đoán và xử lý ticket.

Dựa trên các nghiên cứu trên, chúng tôi xây dựng một hệ thống ITSM tự động hóa bằng cách tích hợp **ServiceNow, Splunk và các mô hình Machine Learning**. Hệ thống vận hành theo các bước sau: **Splunk** thu thập và phân tích log hệ thống bằng các thuật toán **Isolation Forest, LSTM và Autoencoder** để phát hiện bất thường và dự đoán lỗi. Khi xảy ra sự cố, **ServiceNow** sử dụng **NLP (BERT, TF-IDF, SVM)** để phân loại ticket, xác định mức độ ưu tiên và đề xuất phương án khắc phục. Các mô hình **Decision Trees và Reinforcement Learning** giúp hệ thống học từ dữ liệu lịch sử, cải thiện độ chính xác và tối ưu hóa quy trình xử lý. Việc tự động hóa này giúp giảm thời gian phản hồi, hạn chế sai sót và nâng cao hiệu suất vận hành.

MỤC TIÊU

1. Xây dựng hệ thống ITSM tự động hóa bằng cách tích hợp ServiceNow, Splunk và Machine Learning, nhằm tối ưu hóa quy trình giám sát, phát hiện và xử lý sự cố trong hệ thống CNTT.
2. Ứng dụng Machine Learning để phân tích và dự đoán sự cố, sử dụng các thuật toán như Isolation Forest, LSTM, Autoencoder để phát hiện bất thường trong log hệ thống và mô hình NLP (BERT, TF-IDF, SVM) để phân loại ticket, xác định mức độ ưu tiên.
3. Cải thiện độ chính xác và giảm thời gian xử lý ticket bằng cách sử dụng Decision Trees và Reinforcement Learning để tự động đề xuất phương án khắc phục, tối ưu hóa quy trình ITSM và giảm thiểu lỗi do con người.

NỘI DUNG

Nội dung 1: Nghiên cứu tổng quan

- **Mục tiêu:**
    - Hiểu rõ về ITSM, ServiceNow, Splunk và AWS để xác định cách tích hợp hệ thống.
    - Nghiên cứu ứng dụng Machine Learning trong ITSM như giải pháp MyWizard của Accenture, Freshdesk để tối ưu hóa quy trình quản lý sự cố
    - Đánh giá các thuật toán ML phù hợp như Naive Bayes, Random Forest, Gradient Boosting, LSTM, NLP, Reinforcement Learning.
- **Phương pháp:**
    - Tiến hành khảo sát tài liệu khoa học, bài báo và nghiên cứu trước đây liên quan đến Machine Learning trong ITSM.
    - Phân tích cách các hệ thống ITSM hiện tại vận hành và xác định điểm hạn chế khi chưa có tự động hóa.
    - Đánh giá hiệu suất của các thuật toán ML từ các nghiên cứu trước để lựa chọn mô hình phù hợp cho hệ thống.

Nội dung 2: Thiết kế kiến trúc hệ thống

- **Mục tiêu:**
  - Xây dựng kiến trúc tổng thể của hệ thống ITSM tự động hóa.
  - Xác định cách dữ liệu di chuyển giữa Splunk (thu thập log), ServiceNow (quản lý ticket), AWS.
  - Thiết kế cơ chế tích hợp API giữa các hệ thống.
- **Phương pháp:**
  - Xây dựng sơ đồ kiến trúc hệ thống thể hiện mối quan hệ giữa các thành phần.
  - Phân tích luồng dữ liệu, cách thu thập, xử lý và phản hồi thông tin sự cố.
  - Lựa chọn phương thức tích hợp API giữa Splunk, ServiceNow, AWS Lambda để đảm bảo tính chính xác và hiệu suất cao.

Nội dung 3: Xây dựng hệ thống thu thập và phân tích log với Splunk

- **Mục tiêu:**
  - Thu thập log từ hệ thống máy chủ, container và ứng dụng AWS.
  - Áp dụng Machine Learning để phát hiện bất thường và dự đoán lỗi từ dữ liệu log.
  - Xây dựng dashboard Splunk để giám sát dữ liệu theo thời gian thực.
- **Phương pháp:**
  - Kết nối Splunk với các dịch vụ AWS (CloudWatch, EC2, Kubernetes) để thu thập log.
  - Phát triển mô hình phát hiện bất thường sử dụng Isolation Forest, LSTM, Autoencoder.
  - Xây dựng dashboard Splunk để hiển thị dữ liệu giám sát một cách trực quan.

Nội dung 4: Phát triển mô hình Machine Learning để phân loại và dự đoán sự cố

- **Mục tiêu:**
  - Phát triển mô hình Machine Learning giúp phân loại ticket sự cố tự động.
  - Dự đoán lỗi hệ thống và xác định mức độ nghiêm trọng để xử lý kịp thời.
- **Phương pháp:**
  - Thu thập tập dữ liệu từ lịch sử ticket, log sự cố, AWS system metrics.
  - Huấn luyện mô hình NLP (BERT, TF-IDF, SVM) để phân loại ticket.
  - Huấn luyện mô hình XGBoost, Random Forest, Decision Trees để dự đoán lỗi hệ thống và mức độ ưu tiên.
  - So sánh hiệu suất các thuật toán và tinh chỉnh siêu tham số để tối ưu hóa độ chính xác.

Nội dung 5: Tích hợp ServiceNow để tự động xử lý sự cố

- **Mục tiêu:**
  - Xây dựng quy trình tự động tạo, phân loại và xử lý ticket trên ServiceNow.
  - Kết nối ServiceNow với Splunk để nhận cảnh báo sự cố theo thời gian thực.
- **Phương pháp:**

- Cấu hình API giữa ServiceNow và Splunk để tự động tạo ticket khi phát hiện sự cố.
- Xây dựng playbook tự động trên ServiceNow để thực hiện các hành động khắc phục sự cố.

Nội dung 6: Triển khai hệ thống và đánh giá hiệu suất

- **Mục tiêu:**
  - Tích hợp các thành phần thành hệ thống ITSM tự động hóa hoàn chỉnh.
  - Kiểm tra và tối ưu hiệu suất dựa trên thời gian xử lý ticket và độ chính xác của mô hình Machine Learning.
- **Phương pháp:**
  - Kiểm tra toàn bộ hệ thống bằng cách mô phỏng các sự cố thực tế.
  - Đánh giá hiệu suất của mô hình ML trong việc phân loại và dự đoán sự cố.
  - So sánh thời gian xử lý ticket trước và sau khi áp dụng Machine Learning để đo lường hiệu quả.

KẾT QUẢ MONG ĐỢI

- Hệ thống ITSM có thể tự động phát hiện, phân loại và xử lý ticket sự cố dựa trên dữ liệu log, giúp giảm tải công việc cho đội ngũ vận hành IT.
- Cải thiện độ chính xác trong phát hiện bất thường và dự đoán sự cố, giúp giảm thiểu lỗi sai trong quá trình phân tích dữ liệu log thủ công.
- ServiceNow và Splunk được tích hợp để tự động tạo ticket, gán mức độ ưu tiên, và triển khai biện pháp khắc phục như khởi động lại dịch vụ hoặc mở rộng tài nguyên khi cần thiết.

TÀI LIỆU THAM KHẢO

[1] "ServiceNow," Available: https://www.servicenow.com/.

[2] "Splunk," Available: https://www.splunk.com/.

[3] Accenture, "Accenture myWizard," Available: https://www.accenture.com/us-en/services/applied-intelligence/mywizard-intelligent-automation-platform.

[4] "Freshworks," Available: https://www.freshworks.com/vi/freshdesk/compare-helpdesks/.

[5] Zuev, Dmitry & Kalistratov, Alexey & Zuev, Andrey. (2018). Machine Learning in IT Service Management. Procedia Computer Science. 145. 675-679. 10.1016/j.procs.2018.11.063.

[6] R. Gupta, "Automating ITSM Incident Management Process," in *Proceedings of the 2008 International Conference on Autonomic Computing*, Chicago, IL, USA, 2008, pp. 45-52.