|  |  |
| --- | --- |
|  | BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM** |

**ĐỒ ÁN MÔN HỌC**

**TÌM HIỂU KỸ THUẬT WEB SCRAPING BẰNG SELENIUM, THU THẬP VÀ PHÂN TÍCH DỮ LIỆU TỪ TRANG CÁ NHÂN TRÊN MẠNG XÃ HỘI X**

Ngành: **KHOA HỌC DỮ LIỆU**

Môn học: **MÃ NGUỒN MỞ TRONG KHOA HỌC DỮ LIỆU**

Giảng viên hướng dẫn: ThS.Lê Nhật Tùng

Sinh viên thực hiện: 2286400022 - Hồ Nguyễn Hoàng Phát

2286400042 - Trần Lê Vân

TP. Hồ Chí Minh, 2024

|  |  |
| --- | --- |
|  | BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM** |

**ĐỒ ÁN MÔN HỌC**

**TÌM HIỂU KỸ THUẬT WEB SCRAPING BẰNG SELENIUM, THU THẬP VÀ PHÂN TÍCH DỮ LIỆU TỪ TRANG CÁ NHÂN TRÊN MẠNG XÃ HỘI X**

Ngành: **KHOA HỌC DỮ LIỆU**

Môn học: **MÃ NGUỒN MỞ TRONG KHOA HỌC DỮ LIỆU**

Giảng viên hướng dẫn: ThS.Lê Nhật Tùng

Sinh viên thực hiện: 2286400022 - Hồ Nguyễn Hoàng Phát

2286400042 - Trần Lê Vân

TP. Hồ Chí Minh, 2024

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

TPHCM, ngày……tháng……năm 2024

Giáo viên hướng dẫn

(Ký tên, đóng dấu)

**LỜI CAM ĐOAN**

Chúng tôi, Hồ Nguyễn Hoàng Phát, Trần Lê Vân xin cam đoan rằng:

Tất cả thông tin và phân tích trình bày trong báo cáo này được thực hiện một cách chính xác và trung thực.

Mọi dữ liệu, nhận định hoặc ý kiến được trích dẫn từ các nguồn khác đều đã được nêu rõ nguồn gốc và trích dẫn đúng quy định. Tôi cam đoan rằng không có bất kỳ hành vi sao chép hoặc sử dụng thông tin không hợp pháp nào từ các nguồn khác.

Bài báo cáo này là kết quả của công trình nghiên cứu độc lập của chúng tôi và chưa từng được công bố tại bất kỳ nơi nào khác. Tôi cam đoan đã tuân thủ nghiêm ngặt các quy tắc và quy định của môn học, bao gồm việc tham khảo và áp dụng các công cụ nghiên cứu một cách hợp lệ. Nếu phát hiện có bất kỳ sự gian lận nào chúng tôi xin hoàn toàn chịu trách nhiệm về nội dung bài báo cáo của mình.

Tôi hy vọng rằng bài báo cáo này sẽ cung cấp những thông tin hữu ích cho các nhà nghiên cứu , doanh nghiệp. Góp phần vào việc hiểu rõ hơn về mạng xã hội ngày nay.

TPHCM, ngày……tháng 10 năm 2024

**Sinh viên**

**MỤC LỤC**

**DANH MỤC CÁC KÝ HIỆU , CÁC CHỮ VIẾT TẮT**

**DANH MỤC CÁC BẢNG**

**DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ**

# CHƯƠNG 1. TỔNG QUAN

* 1. Giới thiệu đề tài
  2. Những lý thuyết và nghiên cứu trước đây có liên quan
  3. Nhiệm vụ đồ án
     1. Tính cấp thiết và lý do hình thành đề tài
     2. Ý nghĩa khoa học và thực tiễn
     3. Mục tiêu nghiên cứu
     4. Đối tượng và phạm vi giới hạn
  4. Cấu trúc đồ án
     1. Trình bày cấu trúc đồ án
     2. Tóm tắt từng chương

# CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1. Khái niệm và phương pháp giải quyết vấn đề

2.2. Mô tả các công nghệ, hệ thống, các ràng buộc

2.3. Các giải pháp mới và mô hình toán

2.4. Lý giải xây dựng mô hình

# CHƯƠNG 3: PHƯƠNG PHÁP THỰC NGHIỆM

3.1. Phương pháp thu thập dữ liệu:CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM

4.1. Cài đặt Selenium