**TRƯỜNG ĐẠI HỌC SÀI GÒN**

**KHOA CÔNG NGHỆ THÔNG TIN**

-------------------------------------------



**ĐỀ CƯƠNG NGHIÊN CỨU KHOA HỌC**

**Phương Pháp Nghiên Cứu Khoa Học**

**PHÂN TÍCH CẢM XÚC TRONG ĐÁNH GIÁ SẢN PHẨM SỬ DỤNG PHƯƠNG PHÁP HỌC MÁY**

**Giảng Viên Hướng Dẫn : ThS. Đỗ Như Tài**

**Nhóm Sinh Viên Thực Hiện**

**Diệp Thụy An 3122410001**

**Đỗ Mai Anh 3122400006**

**Cao Tiến Cường 3122410043**

**Võ Hoàng Phúc Hy 3123410142**

**Tháng 4/2025 - Thành Phố Hồ Chí Minh**

**MỤC LỤC**

[**ĐỀ CƯƠNG NGHIÊN CỨU KHOA HỌC 4**](#_ltvmvu5jm9ua)

[1. Lý do chọn đề tài 4](#_fznhle5bdf1)

[2. Vấn đề nghiên cứu | hướng tiếp cận 4](#_b86t3gvyzsur)

[3. Mục tiêu và nhiệm vụ nghiên cứu 4](#_unew99m20j)

[4. Đối tượng và phạm vi nghiên cứu 5](#_fiw1mx81kelb)

[5. Phương pháp nghiên cứu 5](#_mf43ylfsvsyr)

[5.1. Phương pháp lý thuyết 5](#_a9yfpczdo7xl)

[5.2. Phương pháp thực nghiệm 6](#_mcerce6hac9v)

[5.3. Phương pháp chuyên gia 6](#_z4tgg815akug)

[6. Dự kiến kế hoạch nghiên cứu 6](#_rx44wzltfsm)

[7. Dự kiến nội dung nghiên cứu 6](#_umm5ezl8et3f)

[8. Danh mục tài liệu tham khảo 6](#_h9xn0jm5wnw0)

**LỜI CẢM ƠN**

Nhóm chúng em xin chân thành cảm ơn **Trường Đại học Sài Gòn** đã tạo điều kiện thuận lợi về cơ sở vật chất và môi trường học tập để chúng em có thể thực hiện và hoàn thành bài báo cáo này.

Đặc biệt, nhóm xin gửi lời cảm ơn sâu sắc đến **TS.Đỗ Như Tài** – giảng viên hướng dẫn, người đã tận tình chỉ bảo, định hướng và hỗ trợ chúng em trong suốt quá trình thực hiện đề tài. Những ý kiến đóng góp quý báu của Thầy là kim chỉ nam giúp nhóm hoàn thiện bài báo cáo một cách tốt nhất.

Nhóm đã nỗ lực hết mình trong quá trình học tập, nhưng do hạn chế về kiến thức và kinh nghiệm, bài báo cáo chắc chắn vẫn còn những thiếu sót. Nhóm rất mong nhận được sự góp ý từ Thầy để hoàn thiện hơn trong những lần sau.

*Nhóm chúng em xin chân thành cảm ơn!*

# 

# **ĐỀ CƯƠNG NGHIÊN CỨU KHOA HỌC**

## **Lý do chọn đề tài**

Trong thời đại số hóa hiện nay, các nền tảng thương mại điện tử đang ngày một phát triển và nắm giữ một vai trò quan trọng trong việc kết nối các doanh nghiệp với khách hàng. Một trong những yếu tố ảnh hưởng trực tiếp đến quyết định mua hàng của khách hàng đó chính là những đánh giá, nhận xét từ những khách hàng trước đó.

Tuy nhiên, số lượng đánh giá được tạo ra mỗi ngày vô cùng lớn khiến việc phân tích thủ công gặp nhiều khó khăn. Do đó việc phân tích cảm xúc (Sentiment Analysis) kết hợp với các phương pháp học máy (Machine Learning), đã trở thành một công cụ mạnh mẽ giúp tự động hóa quá trình phân tích và phân loại cảm xúc của khác hàng. Việc áp dụng học máy không chỉ giúp các doanh nghiệp hiểu rõ hơn về mức độ hài lòng của khách mà còn hỗ trợ việc cải tiến sản phẩm, tạo chiến lược phát triển phù hợp.

## **Vấn đề nghiên cứu**

Việc phân tích cảm xúc của người dùng hiện nay gặp phải nhiều thách thức. Một trong những vấn đề lớn là các mô hình học máy khác nhau sẽ cho ra những kết quả khác nhau, và việc lựa chọn mô hình phù hợp nhất để đạt được hiệu quả là vô cùng quan trọng. Bên cạnh đó, quá trình tiền xử lý văn bản cũng đóng vai trò then chốt trong việc nâng cao chất lượng phân tích. Một bài đánh giá có thể chứa rất nhiều thông tin, trong đó một số thông tin có thể không liên quan đến sản phẩm, gây nhiễu cho quá trình phân tích.

Hướng tiếp cận …

## **Mục tiêu và nhiệm vụ nghiên cứu**

**Mục tiêu nghiên cứu:** Trước thực trạng được nêu ra, mục tiêu của nhóm nghiên cứu là triển khai và đánh giá các mô hình học máy nhằm so sánh hiệu suất và độ chính xác giữa các phương pháp khác nhau. Cụ thể, nhóm sẽ xây dựng và áp dụng một số mô hình học máy như **Logistic Regression** và **Naive Bayes** trên tập dữ liệu đánh giá sản phẩm thực phẩm từ Amazon.

Các mô hình sẽ được đánh giá dựa trên các chỉ số như **Accuracy**, **Precision**,... nhằm xác định mô hình tối ưu nhất. Từ đó, nhóm đề xuất một mô hình phù hợp có thể tích hợp vào các hệ thống thực tế, hỗ trợ doanh nghiệp trong việc phân tích cảm xúc khách hàng, cải thiện sản phẩm và nâng cao trải nghiệm người dùng.

**Nhiệm vụ nghiên cứu:**

## **Đối tượng và phạm vi nghiên cứu**

Trong bài nghiên cứu này nhóm sẽ tập trung vào việc phân tích các đánh giá có trong tập dữ liệu Amazon Fine Food Reviews - một tập đánh giá uy tín , chứa hàng ngàn đánh giá thực phẩm của người dùng.

Nhóm sẽ áp dụng các kĩ thuật xử lí ngôn ngữ tự nhiên (NLP) để khai phá các dữ liệu từ đó so sánh hiệu quả của các mô hình học máy. Đồng thời nhóm cũng áp dụng kĩ thuật trích xuất đặc trưng văn bản để xác định các thông tin quan trọng trong việc đánh giá, góp phần nâng cao độ chính xác cho quá trình phân loại.

Nghiên cứu của nhóm chỉ tập trung vào việc phân loại cảm xúc của dữ liệu đó là tích cực, tiêu cực, trung lập.

## **Phương pháp nghiên cứu**

Bài nghiên cứu này là sự kết hợp chặt chẽ giứa các phương pháp lý thuyết , phương pháp thực nghiệm, phương pháp chuyên gia nhằm đảm bảo tính khoa học, tính nhất quán và hệ thống trong quá trình nghiên cứu đề tài.

### **5.1. Phương pháp lý thuyết**

Tìm hiểu các bài báo có liên quan đền chủ đề phân tích cảm xúc bằng mô hình máy học thông qua các bài nghiên cứu, bài báo trên các nguồn uy tín.

Tổng hợp và phân tích các cơ sở lý thuyết về kĩ thuật xử lí ngôn ngữ tự nhiên (NLP), trích xuất đặc trưng và các khái niêm các mô hình máy học .

Tìm đến các nguồn báo uy tín từ các tạp chí khoa học hàng đầu như IEEE, Springer và các hội nghị trí tuệ nhân tạo (NeurIPS, ACL, EMNLO) để đảm bảo sự uy tín cho bài nghiên cứu.

### **5.2. Phương pháp thực nghiệm**

Phân tích , thiết kế triển kahi các mô hình học máy trên tập dữ liệu Amazon Fine Food Reviews

Cài đặt các mô hình và thuật toán bằng python và sử dụng các thư viện liên quan

Thực hiện huấn luyện và kiểm thử các mô hình, ghi nhận và đánh giá hiệu suất các mô hình bằng độ đo

Thực hiện phân tích so sánh các mô hình và lựa chọn mô hình phù hợp với thực tế nhất

### **5.3. Phương pháp chuyên gia**

Trao đổi với giảng viên hướng dẫn để làm rõ các khuất mắt trong qui trình thực hiện đồng thời tiếp nhận các đánh giá và phản hổi nhằm cải thiện bài nghiên cứu và đưa ra mục tiêu , nhiêm vụ phù hợp cho từng giai đoạn

## **Dự kiến kế hoạch nghiên cứu**

| STT | Nội dung | Dự kiến thời gian thực hiện |
| --- | --- | --- |
|  |  |  |

## **Dự kiến nội dung nghiên cứu**

## **Danh mục tài liệu tham khảo**