**MỤC LỤC**

**[MỞ ĐẦU 1](#_heading=h.1fob9te)**

**[CHƯƠNG 1. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU 4](#_heading=h.3znysh7)**

**[1.1. Tình hình nghiên cứu trong nước 4](#_heading=h.2et92p0)**

**[1.2. Tình hình nghiên cứu nước ngoài 5](#_heading=h.3dy6vkm)**

**[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 11](#_heading=h.26in1rg)**

**[2.1. Tổng quan về mô hình ngôn ngữ lớn LLM 11](#_heading=h.lnxbz9)**

*[2.1.1. Mô hình VGG19 11](#_heading=h.35nkun2)*

*[2.1.2. Mô hình ResNet18 11](#_heading=h.1ksv4uv)*

**[2.2. Tổng quan về mô hình nhận dạng cảm xúc âm nhạc 11](#_heading=h.44sinio)**

*[2.2.1. Mô hình Support Vector Machine (SVM) 11](#_heading=h.2jxsxqh)*

*[2.2.2. Mô hình Linear Discriminat Anlysis (LDA) 12](#_heading=h.z337ya)*

**[CHƯƠNG 3. XỬ LÝ BỘ DỮ LIỆU 13](#_heading=h.3j2qqm3)**

**[3.1. Bộ dữ liệu FER-2013 13](#_heading=h.1y810tw)**

**[3.2. Bộ dữ liệu của Spotify 15](#_heading=h.3whwml4)**

**[CHƯƠNG 4. MÔ HÌNH NHẬN DIỆN CẢM XÚC KHUÔN MẶT VÀ MÔ HÌNH NHẬN DIỆN CẢM XÚC ÂM NHẠC 19](#_heading=h.3o7alnk)**

**[4.1. Mô hình nhận diện cảm xúc khuôn mặt 19](#_heading=h.ihv636)**

*[4.1.1. Mô hình VGG19 19](#_heading=h.32hioqz)*

*[4.1.2. Mô hình ResNet18 20](#_heading=h.41mghml)*

**[4.2. Mô hình nhận diện cảm xúc âm nhạc 21](#_heading=h.3fwokq0)**

**[4.3. Biểu diễn mối quan hệ giữa cảm xúc khuôn mặt và cảm xúc âm nhạc 22](#_heading=h.4f1mdlm)**

**[CHƯƠNG 5. THỰC NGHIỆM VÀ ĐÁNH GIÁ 25](#_heading=h.3tbugp1)**

**[5.1. Môi trường thực nghiệm 25](#_heading=h.28h4qwu)**

**[5.2. Dữ liệu đầu vào 25](#_heading=h.nmf14n)**

**[5.3. Kết quả thực nghiệm 25](#_heading=h.37m2jsg)**

*[5.3.1. Kết quả mô hình nhận diện cảm xúc khuôn mặt trên FER-2013 25](#_heading=h.1mrcu09)*

*[5.3.2. Kết quả mô hình nhận diện cảm xúc âm nhạc trên bộ dữ liệu Spotify 29](#_heading=h.2zbgiuw)*

*[5.3.3. Xây dựng một ứng dụng đề xuất âm nhạc 32](#_heading=h.4bvk7pj)*

**[CHƯƠNG 6. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 34](#_heading=h.2r0uhxc)**

**[6.1. Kết luận 34](#_heading=h.1664s55)**

**[6.2. Hướng phát triển 34](#_heading=h.3q5sasy)**

**LỜI CẢM ƠN**

**DANH MỤC CÁC BẢNG BIỂU**

**DANH MỤC CÁC HÌNH VẼ**

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

##### MỞ ĐẦU

1. **Lý do chọn đề tài**

***1.1. Sự cần thiết nghiên cứu***

- Tin giả xuất hiện rất nhiều và tràn lan trên nhiều trang mạng xã hội, khó kiểm soát.

- Gây ra hiểu lầm, hoang mang dư luận, thậm chí có thể là lợi dung danh tính nhằm mục đích lừa đảo, làm ảnh hưởng tiêu cực đến xã hội.

==> Cần có một hệ thống có khả năng tự động hỗ trợ phát hiện và cảnh báo tin giả.

***1.2. Căn cứ khoa học vấn đề***

- Trước đây đã có nhiều phương pháp sử dụng Machine Learning như Naive Bayes, SVM, LSTM,… nhưng độ chính xác chưa cao.

- Mô hình ngôn ngữ lớn (LLM) có khả năng hiểu ngữ cảnh và xử lý ngôn ngữ tự nhiên tốt hơn.

- Đây là nền tảng có khoa học để áp dụng LLM vào bài toán phát hiện tin tức giả.

1. **Mục tiêu đề tài**    
    - Xây dựng một mô hình ngôn ngữ LLM để có khả năng phát hiện tin tức giả trên không gian mạng.  
    - Xây dựng nên một giao diện web để người dùng có thể dễ dàng truy cập và sử dụng nhằm mục đích phát hiện tin tức giả do người dùng cung cấp.  
    - Giúp mọi người có thể xác định được tin tức thật và giả, từ đó có thể giúp họ biết đâu là tin tức thật nên được tin tưởng và đâu là những tin tức giả không nên được tin tưởng.
2. **Đối tượng nghiên cứu và phạm vi dữ liệu nghiên cứu**

***3.1. Đối tượng nghiên cứu*** *-* Tin tức giả, tin tức sai sự thật, tin tức bị bóp méo nội dung trên không gian mạng Việt Nam.  
 - Các dạng tin giả phổ biến:  
 + Tin tức giật gân, bịa đặt để câu view.  
 + Tin mạo danh tổ chức/cá nhân nhằm lừa đảo.  
 *+* Tin xuyên tạc, bóp méo sự thật nhằm gây hoang mang dư luận.  
 - Ý nghĩa việc chọn đối tượng: phù hợp với thực tế, đảm bảo mô hình phát hiện tin giả có tính ứng dụng cao trong bối cảnh xã hội Việt Nam.

***3.2. Phạm vi dữ liệu nghiên cứu***

- Đa số là những tin tức từ các nguồn mạng xã hội như Facebook, Instagram,…

- Đồng thời cũng có thể lấy nguồn tin tức từ những trang bào chính thống của Việt Nam như VNexpress, tuoitre,…

- Ngôn ngữ nghiên cứu và kiểu dữ liệu: tập trung hết vào những dữ liệu dạng văn bản được viết bằng tiếng Việt.

- Giới hạn nghiên cứu: chỉ tập trung vào tin tức giả có liên quan đến các vấn đề xã hội, chính trị, y tế, giải trí… không bao gồm tin nhắn riêng tư hoặc dữ liệu cá nhân.

1. **Phương pháp nghiên cứu** 
   1. ***Phương pháp nghiên cứu lý thuyết*** - Thu thập và tổng hợp các tài liệu, công trình nghiên cứu trong và ngoài nước liên quan đến bài toán phát hiện tin tức giả, từ đó xây dựng cơ sở lý luận cho đề tài.  
       - Phân tích đặc điểm của mô hình ngôn ngữ lớn (LLM), cơ chế của Transformer để làm nền tảng khoa học cho việc áp dụng vào bài toán phân loại.
   2. ***Phương pháp nghiên cứu thực nghiệm***    
       - Xây dựng và xử lý bộ dữ liệu gồm những tin tức thật và tin tức giả, đảm bảo sự chất lượng cho giai đoạn huấn luyện mô hình.  
       - Thực hiện fine-tuning trên mô hình LLM đã huấn luyện từ trước, điều chỉnh siêu tham số và đánh giá hiệu quả bằng các chỉ số như Accuracy, F1-Score, Brier-score, ECE.
2. **Bố cục báo cáo**

Báo cáo này gồm 06 chương:

Chương 1: Tổng quan tình hình nghiên cứu.

Chương 2: Cơ sở lý thuyết.

Chương 3: Xây dựng bộ dữ liệu.

Chương 4: Mô hình huấn luyện.

Chương 5: Thực nghiệm và đánh giá.

Chương 6: Kết luận và hướng phát triển.

**Chương 1. Tổng quan tình hình nghiên cứu**