**SỞ GIÁO DỤC & ĐÀO TẠO THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG THPT CHUYÊN TRẦN ĐẠI NGHĨA**

**CUỘC THI “HỌC SINH NGHIÊN CỨU KHOA HỌC”**

**NĂM HỌC 2025 - 2026**

**BÁO CÁO NGHIÊN CỨU**

**TÊN DỰ ÁN**

**Học sinh: Nguyễn Long Lớp:8A14**

**Học sinh: Nguyễn Nam Long Lớp:8A8**

**Họ và tên GVHD: Trần Thanh Ý**

**Lĩnh vực: Hệ thống phần mềm**

**Thành phố Hồ Chí Minh, tháng 9 năm 2025**

**MỤC LỤC**

[**TÓM TẮT DỰ ÁN 3**](#_heading=h.l7ncg9rur5z4)

[**NỘI DUNG 4**](#_heading=h.rx38iru2mdxh)

[Chương 1: Lý do chọn đề tài 4](#_heading=h.1ik4whic3jsw)

[Chương 2: Đặt vấn đề nghiên cứu 4](#_heading=h.48916jecd3e1)

[Chương 3: Thiết kế và phương pháp nghiên cứu 4](#_heading=h.flcpvitpkav6)

[Chương 4: Tiến hành nghiên cứu 4](#_heading=h.gluc7fj9bpj6)

[Chương 5: Kết luận khoa học 4](#_heading=h.aijfisk5wnr2)

[**NGUỒN THAM KHẢO 5**](#_heading=h.z5puv5xaqyee)

## TÓM TẮT DỰ ÁN

1. **Tính mới  
   -** Đề tài sử dụng mô hình ngôn ngữ lớn (LLM) để phát hiện tin giả, thay vì các cách truyền thống như Naive Bayes hay SVM. Đây là hướng đi mới, phù hợp với xu thế AI hiện nay.
2. **Tính khoa học**

* Nghiên cứu dựa trên các kiến thức về xử lý ngôn ngữ tự nhiên và học máy. Thử nghiệm với mô hình LLM giúp kiểm chứng khả năng phân tích ngữ cảnh và so sánh với các mô hình cũ.

1. **Tính thực tiễn**

* Tin giả lan nhanh trên mạng xã hội, gây nhiều ảnh hưởng xấu. Hệ thống phát hiện tin giả có thể hỗ trợ người dùng kiểm chứng thông tin và giảm thiểu tác hại.

1. **Tính cộng đồng**

* Sản phẩm có thể phát triển thành công cụ phục vụ cộng đồng, giúp mọi người tiếp cận thông tin chính xác hơn.

## NỘI DUNG

### Chương 1: Lý do chọn đề tài

**1.1 Sự cần thiết nghiên cứu**

* Tin giả xuất hiện nhiều trên mạng xã hội, khó kiểm soát.
* Gây ra hiểu lầm, hoang mang dư luận, ảnh hưởng tiêu cực đến xã hội.
* Cần có hệ thống 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 khoa học để áp dụng LLM vào bài toán phát hiện tin giả.

### Chương 2: Đặt vấn đề nghiên cứu

**2.1 Vấn đề nghiên cứu**

* Tin giả ngày càng lan rộng, khó phân biệt được với tin thật.
* Cần có một hệ thống AI có khả năng tự động nhận diện, thay vì chỉ dựa vào con người kiểm chứng.

**2.2 Câu hỏi nghiên cứu**

* Mô hình LLM có thể phát hiện tin giả hiệu quả và tốt hơn các mô hình truyền thống không?
* Độ chính xác của mô hình LLM trong phân loại tin tức thật/giả đạt mức nào?
* Có thể xây dựng ứng dụng kiểm chứng tin tức cho cộng đồng từ kết quả nghiên cứu không?

**2.3 Các giả thuyết khoa học cần được kiểm định**

* LLM có thể phân loại tin giả với độ chính xác cao.
* LLM cho kết quả tốt hơn các mô hình cũ (Naive Bayes, SVM, LSTM).
* Mô hình có thể ứng dụng vào thực tế để phục vụ cộng đồng.

**2.4 Mục tiêu của nghiên cứu**

**2.4.1 Mục tiêu chung**

Xây dựng hệ thống phát hiện tin giả dựa trên LLM.

**2.4.2 Mục tiêu cụ thể**

* Thu thập và xử lý dữ liệu tin tức thật/giả.
* Huấn luyện và thử nghiệm mô hình LLM.
* So sánh với các phương pháp truyền thống.
* Đề xuất hướng ứng dụng kết quả nghiên cứu.

### Chương 3: Thiết kế và phương pháp nghiên cứu

**3.1 Phương pháp nghiên cứu**

**3.1.1 Nghiên cứu tài liệu**

* Tìm hiểu các công trình trước về phát hiện tin giả.
* Tìm hiểu cách ứng dụng LLM vào trong phân loại văn bản.

**3.1.2 Phương pháp thu thập số liệu**

* Thu thập dataset có sẵn trên Kaggle.

**3.2 Thiết kế nghiên cứu**

### Chương 4: Tiến hành nghiên cứu

### Chương 5: Kết luận khoa học

## NGUỒN THAM KHẢO