| ĐẠI HỌC ĐÀ NẴNG  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG VIỆT - HÀN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc**    *Đà Nẵng, ngày 13 tháng 09 năm 2024* |
| --- | --- |

**ĐỀ XUẤT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN NĂM HỌC 2024-2025**

**1. Thông tin người đề xuất:**

| **- Họ và tên:** Thái Lý Hân | **Lớp:** 22AD | **Mã số SV:** 22AD014 |
| --- | --- | --- |
| **- Email:** hantl.22ds@vku.udn.vn | **Số điện thoại:** 0328612168 | |
| **- Họ và tên:** Phan Đức Huy | **Lớp:** 22AD | **Mã số SV:** 22AD021 |
| **- Email:** huypd.22ds@vku.udn.vn | **Số điện thoại:** 0862600454 | |
| **- Họ và tên:** Nguyễn Đức Gia Phúc | **Lớp:** 22AD | **Mã số SV:** 22AD037 |
| **- Email:** phucndg.22ds@vku.udn.vn | **Số điện thoại:** 0797826527 | |

**2.** **Tên đề tài**: Phát hiện meme có chứa nội dung gây hận thù

**3. Lĩnh vực nghiên cứu**: Trí tuệ nhân tạo

**4. Tính cấp thiết**:

- Meme (hình ảnh châm biếm) là một hình thức truyền tải thông tin.

- Khi mạng xã hội phát triển và ngày càng nhiều người tiếp cận nội dung giải trí trên các nền tảng này, meme trở thành phương tiện truyền tải thông tin phổ biến. Vì meme có thể lan truyền nhanh chóng qua các mạng xã hội và thường có tính chất hài hước, dễ tiếp cận, việc kiểm soát và phát hiện meme có hại trở nên quan trọng để giảm thiểu tác động tiêu cực đến xã hội.

**5. Mục tiêu**: một mô hình AI có khả năng phân tích và nhận diện các meme có nội dung độc hại, ví dụ như thông tin sai lệch, thù hận, phân biệt chủng tộc, hay cổ xúy bạo lực.

**6. Nội dung chính**:

- Thu thập dữ liệu meme, xử lý và phân tích dữ liệu, xác định ngữ cảnh, và phân loại mức độ gây hại của meme.

- Khảo sát, nghiên cứu các phương pháp huấn luyện mô hình AI sử dụng dữ liệu lớn, giúp mô hình học cách nhận diện meme dựa trên các đặc điểm như hình ảnh, văn bản, ngữ cảnh.

- Phát triển mô hình AI phổ biến có khả năng phát hiện nội dung có hại, bao gồm xử lý ngôn ngữ tự nhiên (NLP) và nhận diện hình ảnh (Computer Vision).

**7. Sản phẩm và kết quả dự kiến**:

- Sản phẩm: Một mô hình AI hoàn thiện có khả năng tự động nhận diện và phân loại meme có hại dựa trên các tiêu chí như ngôn ngữ, hình ảnh và ngữ cảnh.

- Kết quả dự kiến: Mô hình sẽ có khả năng phân tích meme trên nhiều nền tảng mạng xã hội, phát hiện các dạng nội dung độc hại như thù hận, phân biệt chủng tộc, bạo lực, hoặc thông tin sai lệch với độ chính xác cao.

**8. Thời gian nghiên cứu dự kiến**: 6 tháng

**Xác nhận của khoa Người hướng dẫn Sinh viên**

(chữ ký, họ và tên) (chữ ký, họ và tên) (chữ ký, họ và tên)