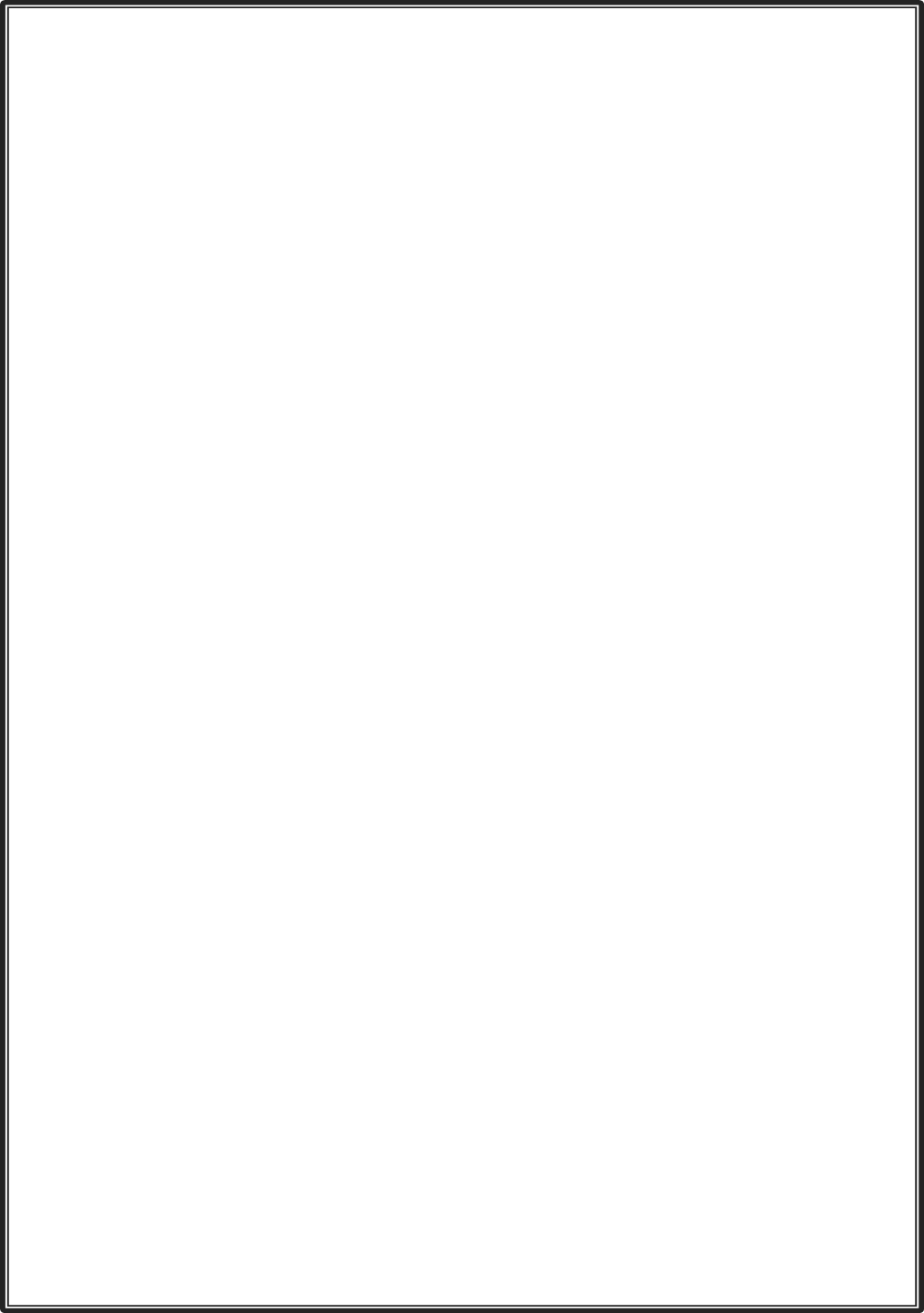
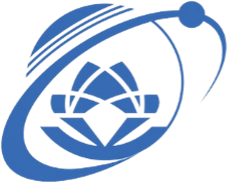
**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA KHOA HỌC MÁY TÍNH**

****

**BÀI TẬP CHƯƠNG 1**

**NHÓM 4**

**MÔN HỌC: KHAI THÁC DỮ LIỆU VÀ ỨNG DỤNG**

**GVHD: NGUYỄN THỊ ANH THƯ**

**TP. Hồ Chí Minh, 2/2024**

**1. Bài toán được dùng trong vấn đề này là bài toán phân lớp nhóm tính cách**.

Cụ thể mỗi người dùng Facebook sẽ có 5 tiêu chí tính cách khác nhau (được áp dụng từ 5 bài thử nghiệm tính cách tổng quát trong tâm lý học) và mức độ tương ứng của chúng, gồm: Độ cởi mở, Sự tận tâm, Hướng ngoại, Tính dễ chịu và Tâm lý bất ổn. Các thông tin này đều được lấy từ các dữ liệu hoạt động tương tác trên Facebook của họ (chẳng hạn thích bài viết nào, kết bạn với ai, …). Sau đó, công ty Cambridge Analytica sẽ sử dụng thông tin này và phân loại mỗi người dùng vào một nhóm tính cách (ví dụ người Mạo hiểm sẽ có cả 5 điểm số đều cao). Cuối cùng thông tin này sẽ được Facebook sử dụng để đề xuất các quảng cáo phù hợp với từng kiểu người và chính sách của tổng thống tương ứng mà tổ chức Cambridge Analytica ủng hộ nhằm tăng khả năng họ sẽ đồng tình và bỏ phiếu cho vị tổng thống.

**2. Kỹ thuật IT được sử dụng:**

* Kỹ thuật thu thập dữ liệu: Để lấy thông tin từ các nguồn khác nhau như Facebook và các trang web khác.
* Quy trình tiếp cận thông tin (Information Retrieval): Để truy xuất và xử lý thông tin có ích từ dữ liệu thu thập được.
* Phân tích dữ liệu, chuyển đổi data về dạng dữ liệu có cấu trúc với 5 đặc trưng (Độ cởi mở, Sự tận tâm, Hướng ngoại, Tính dễ chịu và Tâm lý bất ổn) với khoảng giá trị từ 0-100%: Để xử lý, phân tích và làm sạch lượng dữ liệu để phục vụ các mục đích khác nhau.
* Machine Learning và Data Science: Để áp dụng và phát triển các mô hình dự đoán và phân loại dựa trên dữ liệu đã xử lý.

Kiến thức của các lĩnh vực liên quan: Tâm lý học.

**3. Kết quả của bài toán:** là nhóm tính cách của người dùng (thuộc 1 trong 3 nhóm tính cách sau: người mạo hiểm, người bảo vệ, người lãnh đạo)

**4. Các chuyên gia đã vận dụng kết quả vào bầu cử tổng thống như sau:** Dựa vào từng tính cách của một người thì sẽ chạy quảng cáo chính trị (sử dụng súng) tương ứng.

Với kết quả được đánh giá dựa trên 5 đặc trưng tính cách (Tính hướng ngoại, Tính tận tâm, Tính cởi mở với trải nghiệm, Tính tương hợp, Tính bất ổn cảm xúc), các chuyên gia phân loại thành từng nhóm tính cách đối tượng và sau đó sẽ đưa những quảng cáo truyền bá thông điệp tư tưởng và định hướng chính trị của chính trị gia phù hợp đến những nhóm đối tượng dựa trên nhóm tính cách đã được phân loại trên. Các quảng cáo sẽ được tạo ra theo từng nhóm sao cho phù hợp với tính cách, đặc điểm và được liên tục đăng lên trên newsfeed của những người trưởng thành có quyền bầu cử tuỳ thuộc vào nhóm người nhằm hướng sự quan tâm của họ đến định hướng chính trị của ứng viên tổng thống góp phần làm tăng sự ủng hộ của người bầu cử với ứng viên tổng thống.

Vì vậy, do các quảng cáo phù hợp được đưa ra với từng tính cách đã được phân loại sẽ khiến cho người bầu cử thay đổi ý định và làm thay đổi kết quả của cuộc bầu cử tổng thống