**Câu hỏi: Sau khi xem đoạn video về Cambridge Analytica Scandal, thì ứng dụng của Data Mining vào MXH ở đây là classification hay cluster?**

Sau khi nhóm tiến hành thảo luận, nhóm em đã thống nhất đây là bài toán Classification.

Lý giải cho kết luận của nhóm, thì đầu tiên ta sẽ đi vào quy trình khai thác dữ liệu của công ty Cambridge Analytica:

+ Đầu tiên, họ tiến hành đưa ra một bài test tính cách dựa vào các câu trả lời khác nhau thông qua một ứng dụng có tên myPersonality được nghiên cứu bởi các nhà tâm lý học và những người thực hiện sẽ được trả ít nhất là 5$. Kết quả từ bài khảo sát này sẽ được chia làm 5 nhóm người: *“These are: openness (how open you are to new experiences?), conscientiousness (how much of a perfectionist are you?), extroversion (how sociable are you?), agreeableness (how considerate and cooperative you are?) and neuroticism (are you easily upset?).*

+ Tiếp theo, để có thể thực hiện bài test này, người dùng phải đăng nhập bằng tài khoản Facebook, từ đó sẽ giúp cho công ty có được các dữ liệu quan trọng như: số Like, các bài post, hình ảnh, bạn bè,….

+ Dữ liệu thu thập từ facebook và bài khảo sát sẽ được kết hợp lại tạo thành một điểm dữ liệu góp phần vào việc xây dựng mô hình dự đoán, giúp cho chúng ta có thể thao túng mạng xã hội bằng cách đưa ra các thông tin, quảng cáo có ảnh hưởng đến quyết định của người dùng dựa trên nhóm tính cách của họ

**Data:** Dữ liệu từ mạng xã hội như hình ảnh, post, video, bạn bè,…

**Nhãn:** 5 nhóm tính cách đã được người dùng cung cấp thông qua ứng dụng

**Minh, Phước, Thoại, Nguyên, Dũng đóng góp ý kiến**

**Dũng tổng hợp ý kiến**