

## Kịch bản

### 1. Thu thập dữ liệu trên mạng

Để thu thập dữ liệu mạng có thể sử dụng phương pháp Web Crawler. Web Crawler thu thập toàn bộ nội dung của facebook, instagram, twitter gồm các trường thông tin như: nội dung bài viết, người đăng, thời gian đăng bài, ....

Mô hình Crawler như sau:

- Tự động mở trình duyệt Chrome, login vào FACEBOOK.
- Vào một Fanpage bất kỳ, crawl id bài viết trên page.
- Vào từng bài viết dựa vào id đã crawl được.
- Tiến hành Crawl nội dung, comment của bài viết.

Sau khi đã thu thập dữ liệu mạng sử dụng phương pháp Web Crawler. Dữ liệu sẽ được lưu trữ dưới dạng CSV (định dạng CSV sẽ thuận lợi hơn cho những người phân tích và xử lý dữ liệu).

### 2. Xác định nhãn sắc thái bình luận

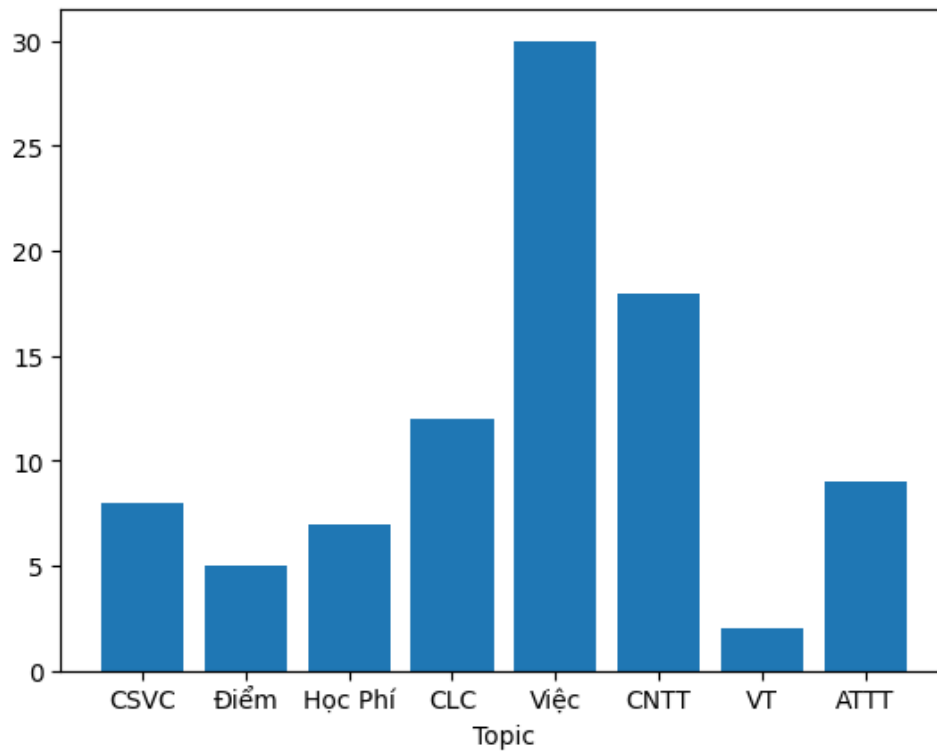
Dữ liệu sau khi được thu thập sẽ sử dụng model AI để gán nhãn comment. Ở đây sẽ có 3 trạng thái comment: positive: 0, neutral: 1, negative: 2.

Comment	Sentiment
Học phí thì đắt, chất lượng giáo viên thì tệ	Negative
Hiện tại Học viện không tuyển sinh văn bằng 2 nhé	Neutral
Trường nào cũng có mặt mạnh, mặt yếu riêng, tùy theo sở thích và năng lực của mình thôi. Khuyến đàn em trước khi vô phải suy nghĩ cho kĩ nha, coi chừng vỡ mộng. Còn riêng đàn anh thì dù 2 năm gần bó, vẫn quyết định bỏ trường mà đi...	Negative
Trường đào tạo tốt, thầy cô giáo nhiệt tình	Positive

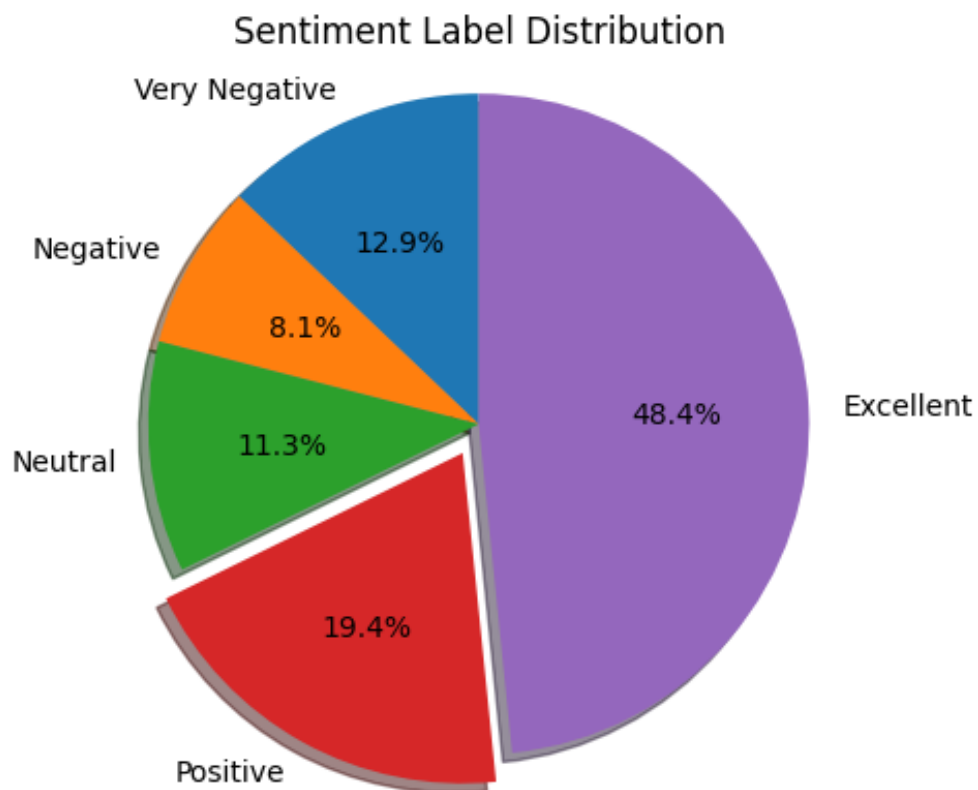
### 3. Phân tích

Phân tích dữ liệu thu thập:

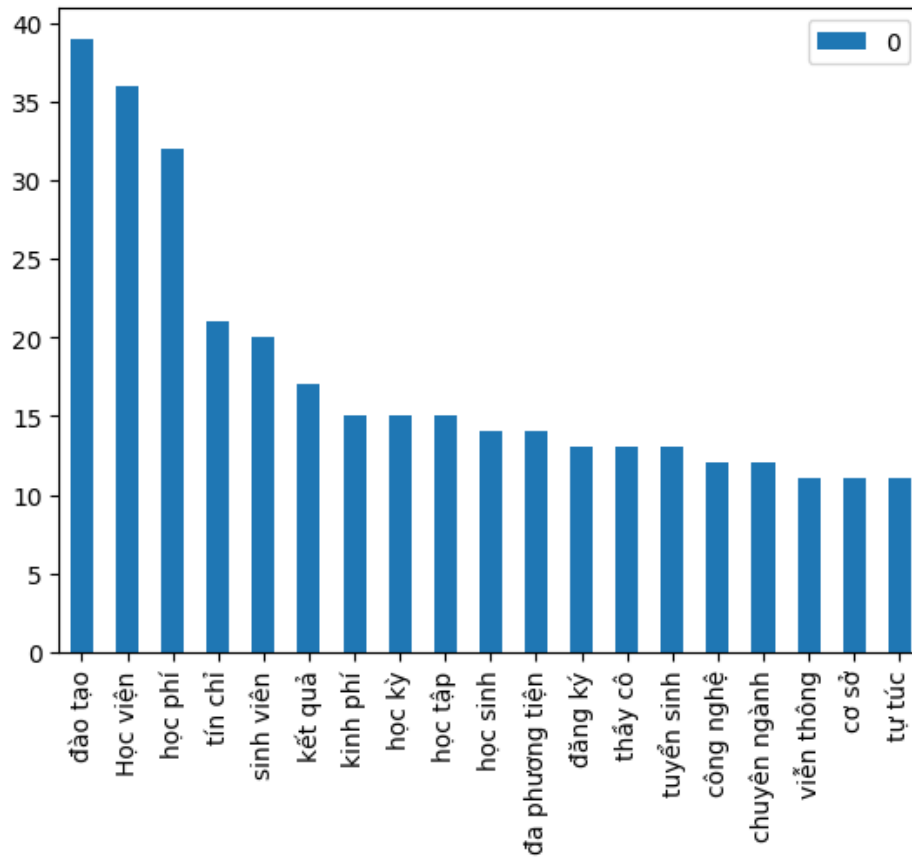
- Chủ đề nào được nhắc đến nhiều nhất



➤ Phân phối dữ liệu nhãn cảm xúc

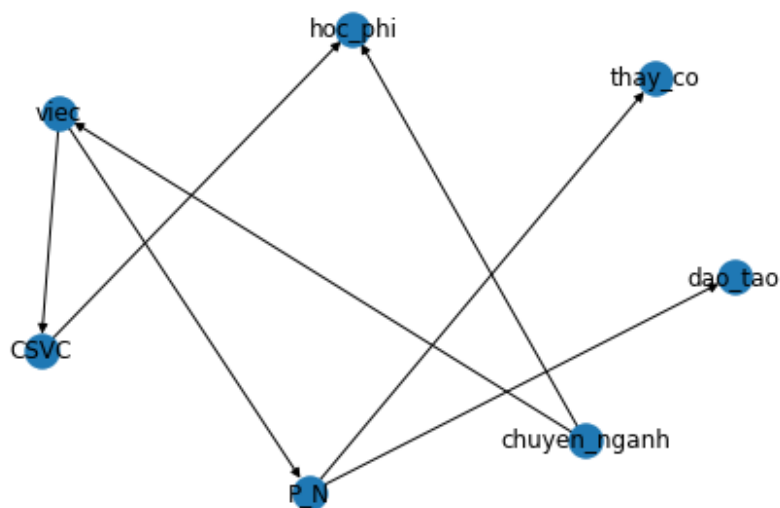


➤ Tần suất xuất hiện của các từ



#### 4. Xây dựng graph model

Dựa trên bảng thống kê tần suất xuất hiện từ chúng ta tiến hành chọn ra các nhân tố quan trọng nhất (thường từ 5-6 nhân tố) để xây dựng đồ thị quan hệ.



## 5. Đánh giá độ ảnh hưởng các nhân tố

- ❖ Đánh giá điểm số cho nhân tố dựa trên 3 cấp độ (trung bình, khá, giỏi)

<b>Đào tạo</b>	<b>Điểm số</b>
Trung bình	0.4286
Khá	0.3571
Giỏi	0.2143

- ❖ Xác định độ ảnh hưởng giữa các nhân tố:

- Chất lượng của giáo viên tốt sẽ ảnh hưởng như nào đến đào tạo:

<b>Đào tạo/Giáo viên tốt</b>	<b>Điểm số</b>
Trung bình	0.0000
Khá	0.6667
Giỏi	0.3333

- Chất lượng của giáo viên tệ sẽ ảnh hưởng như nào đến đào tạo:

<b>Đào tạo/Giáo viên tệ</b>	<b>Điểm số</b>
Trung bình	0.6000
Khá	0.2333
Giỏi	0.1667