

# Báo cáo môn học - Phân tích mạng xã hội

Hạn nộp: 14/10/2023

14th September 2023

## Abstract

Tóm tắt - Bên dưới là hướng dẫn viết báo cáo môn học “Phân tích mạng xã hội”. Hướng dẫn bên dưới có thể không phù hợp với một vài đề tài (nếu có), tuy nhiên, nó được xem như là hướng dẫn “cơ bản/chung” để các bạn viết báo cáo. **Độ dài báo cáo tính theo trang (page) là từ 15 đến 25 trang, khổ giấy A4, font size:12pt.** Nội dung nộp báo cáo: (1) Báo cáo (file MS Word) (2) Slide báo cáo (file MS Powerpoint) (3) Mã nguồn gồm: code, dataset. Nếu dataset và code có kích thước lớn ( $\geq 100$ MB), gửi đường link Google drive đến dataset và code, và cấp quyền truy cập là view. Mọi thắc mắc khác các bạn email đến lamnn@tvu.edu.vn. **Cấu trúc của báo cáo có thể trình bày dưới dạng một bài báo khoa học - tham khảo các bài báo sử dụng trong Báo cáo quá trình.** Để thực hiện báo cáo có thể làm theo cách sau: sử dụng Python (Google Colab) để load dataset lên DataFrame, tiền xử lý dữ liệu, thống kê mô tả, vẽ hình,... Sau đó, xác định cột nào là source, cột nào là target, có cần weight hay không? Tạo DataFrame dạng có các cột source, target, weight (nếu có). Giá trị của từng dòng là node->node, ví dụ: user1|user2|2: link từ user1 -> user2, trọng số link (cạnh) là 2. Lưu DataFrame này vào file csv (sử dụng hàm của Python). Sử dụng Gephi để phân tích cấu trúc mạng, tính các network measurements,... bằng cách import file csv này.

Báo cáo bao gồm các nội dung cơ bản sau (hoặc nhiều hơn):

Link tham khảo [Các link ở mục Projects from Fall 2009](#)

## 1 Giới thiệu

- Lý do chọn đề tài
- Mô tả đề tài (bài toán)
- Liệt kê các mục tiêu nghiên cứu của đề tài (có thể liệt kê dưới dạng: Câu hỏi nghiên cứu (research questions): ví dụ: 1) Bộ phim nào thu hút được người xem/bình luận nhất. 2) Ai là người tham gia tích cực nhất (bình luận, post bài,...). 3) Tác giả nào có nhiều phim nhất, diễn viên nào đóng film nhiều nhất, ... 4) Đặc điểm của cấu trúc mạng (tương tác giữa những người tham gia vào mạng (network). 6) ... 7)...
- Mô tả cách giải quyết đề tài (bài toán)/cách giải đáp các câu hỏi nghiên cứu

## 2 Cơ sở lý thuyết

- Cơ sở lý thuyết liên quan đến đề tài
- Các bài báo, công trình nghiên cứu liên quan đến nội dung/phương pháp nghiên cứu sử dụng trong đề tài
- Phương pháp nghiên cứu đề xuất giống, khác, cải thiện gì so với các công trình nghiên cứu hiện có

## 3 Dataset- dữ liệu

- Mô tả dataset: dataset được lấy ở đâu? Mô tả chi tiết về dataset
- Thống kê dataset (dạng Table)

## 4 Phân tích mô tả

Phân tích mô tả dữ liệu đã thu thập

- Số lượng bình luận, số lượng bình luận theo ngày, ...
- Số lượng bài đã post, số lượng bài post theo ngày, tháng, ...
- Thời gian post/thời gian bình luận...
- Sắc thái của bình luận (tích cực, tiêu cực, trung lập)
- Wordcloud của nội dung (text) câu bình luận, bài post,...
- ...

## 5 Phân tích mạng - Social Network Analysis

- Mô tả cấu trúc mạng - cái gì là node (nút)? có link từ node đến node khi nào?
- Trình bày các network measurement được sử dụng (degree, indegree, ..., pagerank, ...)
- Phân tích dựa trên các network measurements trình bày ở trên: ai là người ảnh hưởng nhất, top những người ảnh hưởng, ...
- Những người ảnh hưởng họ làm gì? hành động của họ (activity - bình luận hay post bài, ....)
- ...

## 6 Phân tích Cộng đồng - Community Analysis

- Trình bày giải thuật sử dụng để phát hiện cộng đồng (community)
- Kết quả phân tích: có các cộng đồng nào?
- Đặc điểm của cộng đồng đó là gì?
- Ai là người hoạt động tích cực trong từng cộng đồng (dựa theo degree, indegree, outdegree, pageRank/hub/authority...)
- ...

## 7 Thảo luận

Thảo luận về đề tài, kết quả phân tích của đề tài.

## 8 Kết luận

Tổng kết ngắn gọn về kết quả đạt được của đề tài, các kết quả quan trọng. Nếu đề tài tiếp tục phát triển, hướng phát triển là gì, nêu các hướng giải quyết vấn đề trong tương lai.

## 9 Tài liệu tham khảo

Tài liệu tham khảo sử dụng trong đề tài, trích dẫn đúng qui định