

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH
KHOA ĐÀO TẠO CHẤT LƯỢNG CAO
NGÀNH CÔNG NGHỆ THÔNG TIN



BÁO CÁO TIẾN ĐỘ -TUẦN 02
MÔN: ĐỒ ÁN 3
ĐỀ TÀI: TÌM HIỂU VỀ THUẬT TOÁN
RECOMMENDATION

GVHD : Thầy Huỳnh Xuân Phụng

SVTH :

Nguyễn Thành Như

17110202

Võ Ngọc Thuận

17110234

TP. Hồ Chí Minh, tháng 10 năm 2020

1. Hiệu quả thuật toán Recommendation đem lại. [Tự tìm hiểu]

- Thuật toán giúp tăng doanh thu của website.
- Gợi ý khách hàng hàng ngày giúp tăng khả năng mua hàng đối với sản phẩm trong tương lai.
- Giúp khách hàng tiếp cận được món hàng cần thiết một cách dễ dàng.
- Giúp sản phẩm đến đúng với khách hàng đang có nhu cầu.

2. Tìm hiểu quy trình, các bước để xây dựng một hệ thống Recommendation. [Đọc từ nhiều nguồn tài liệu]

Các bước để xây dựng một hệ thống Recommendation là:

- **Thu thập là lưu trữ thông tin:** như chúng ta đã biết thì Recommendation là một thuật toán áp dụng Machine Learning, thì thứ đầu tiên phải có đó chính là việc thu thập dữ liệu để tiến hành các bước sau như là đánh giá, phân tích. Những dữ liệu chúng ta tiến hành thu thập đó là các hành động/tương tác của người dùng trên hệ thống đối với một đối tượng/sản phẩm, chẳng hạn như: mua, số lượt xem, đánh giá,... **[Đã hiểu]**
- **Phân tích/chuẩn hóa dữ liệu:** theo nhóm tìm hiểu được thì sau khi hoàn thành bước thu thập thông tin, dữ liệu thì chúng được lưu trữ lại trong database, tuy nhiên lúc này những thông tin chúng ta thu thập được sẽ là một ma trận thưa (trống nhiều vị trí, ví dụ một người A có thể đã xem sản phẩm A, E, F nhưng họ không xem các sản phẩm B, C, D thì lúc này dữ liệu ở người A với sản phẩm B, C, D không có). Chính vì lý do đó nên trước khi đi tính toán độ tương tự của các user, chúng ta cần phải điền đầy đủ các dấu hỏi đó mà nó **không làm ảnh hưởng đến độ tương tự** của ma trận. Sau khi điền đầy đủ thông tin, dữ liệu thì chúng ta mới có thể áp dụng được các thuật toán để tính toán độ tương tự vào. Các cách chuẩn hóa sau đây trong phương pháp User Collaborative Filtering (Lọc cộng tác) :
 - Điền giá trị 0 vào ô trống (chưa tối ưu).
 - Điền giá trị trung bình trong khoảng rating (ví dụ 1-5 thì ta điền 2.5).
 - Lấy trung bình của các rating user khác. Như hình bên dưới **[Chưa hiểu cách ra kết quả cuối cùng ở phần c trong hình bên dưới]**

	u_0	u_1	u_2	u_3	u_4	u_5	u_6
i_0	5	5	2	0	1	?	?
i_1	4	?	?	0	?	2	?
i_2	?	4	1	?	?	1	1
i_3	2	2	3	4	4	?	4
i_4	2	0	4	?	?	?	5
	↓	↓	↓	↓	↓	↓	↓
u_j	3.25	2.75	2.5	1.33	2.5	1.5	3.33

a) Original utility matrix Y and mean user ratings.

	u_0	u_1	u_2	u_3	u_4	u_5	u_6
i_0	1.75	2.25	-0.5	-1.33	-1.5	0	0
i_1	0.75	0	0	-1.33	0	0.5	0
i_2	0	1.25	-1.5	0	0	-0.5	-2.33
i_3	-1.25	-0.75	0.5	2.67	1.5	0	0.67
i_4	-1.25	-2.75	1.5	0	0	0	1.67

b) Normalized utility matrix \bar{Y} .

	u_0	u_1	u_2	u_3	u_4	u_5	u_6
u_0	1	0.83	-0.58	-0.79	-0.82	0.2	-0.38
u_1	0.83	1	-0.87	-0.40	-0.55	-0.23	-0.71
u_2	-0.58	-0.87	1	0.27	0.32	0.47	0.96
u_3	-0.79	-0.40	0.27	1	0.87	-0.29	0.18
u_4	-0.82	-0.55	0.32	0.87	1	0	0.16
u_5	0.2	-0.23	0.47	-0.29	0	1	0.56
u_6	-0.38	-0.71	0.96	0.18	0.16	0.56	1

c) User similarity matrix S .

- **Lọc và đưa ra gợi ý phù hợp:** chọn được mô hình phù hợp từ những dữ liệu đã được chuẩn hóa, chúng ta tiến hành chạy mô hình rồi phân tích để đưa ra gợi ý phù hợp nhất cho người dùng [Nhóm chưa tìm hiểu được cụ thể/chi tiết cách tiến hành theo từng mô hình]

3. Các phương pháp trong thuật toán gợi ý Recommendation.

Các phương pháp trong thuật toán Recommendation:

- Content - Based Recommendation System (Gợi ý dựa vào nội dung): ví dụ như khi ta xem một cuốn sách có thể loại là trinh thám, thì hệ thống sẽ gợi ý ra những sản phẩm có nội dung gần giống vậy, chung thể loại. [Đã hiểu cơ bản, chưa tìm hiểu sâu vào chi tiết]
- Clustering User (Phân cụm người dùng): phân cụm và gợi ý người dùng theo nhóm. [Đã hiểu cơ bản, chưa tìm hiểu sâu vào chi tiết]
- Collaborative Filtering (Lọc cộng tác): đánh giá toàn bộ người dùng và toàn bộ sản phẩm rồi phân tích đưa ra kết quả. Nhóm tìm hiểu thấy thuật toán này nên sử dụng nhất vì tính logic cao nên cho độ chính xác cao hơn. [Đã hiểu cơ bản, chưa tìm hiểu sâu vào chi tiết]
- Session - Based Recommendation System (Gợi ý theo chuỗi hoạt động): dựa vào chuỗi hoạt động của một người dùng để đưa ra gợi ý. Ví dụ: trên trang Tiki, khi click mua/xem một món hàng nào đó, ở phía dưới sẽ hiển thị những sản phẩm thường được mua cùng kèm theo. [Đã hiểu cơ bản, chưa tìm hiểu sâu vào chi tiết]

4. Công việc dự kiến tuần sau:

Tìm hiểu chi tiết hơn về các phương pháp trong thuật toán gợi ý Recommendation.