

# HỆ THỐNG GỢI Ý SẢN PHẨM DỰA TRÊN MÔ HÌNH HỌC NHÂN QUẢ

(Causal learning models for building recommendation system)

**Giáo viên hướng dẫn**

Ths. Trần Trung Kiên  
TS. Nguyễn Ngọc Thảo

**Sinh viên thực hiện**

Nguyễn Duy Vũ  
Nguyễn Chiêu Bản

# Nội dung báo cáo

01

Giới thiệu về đề tài

02

Các lý thuyết nền tảng

03

Phương pháp

04

Kết quả thí nghiệm

01

# Giới thiệu

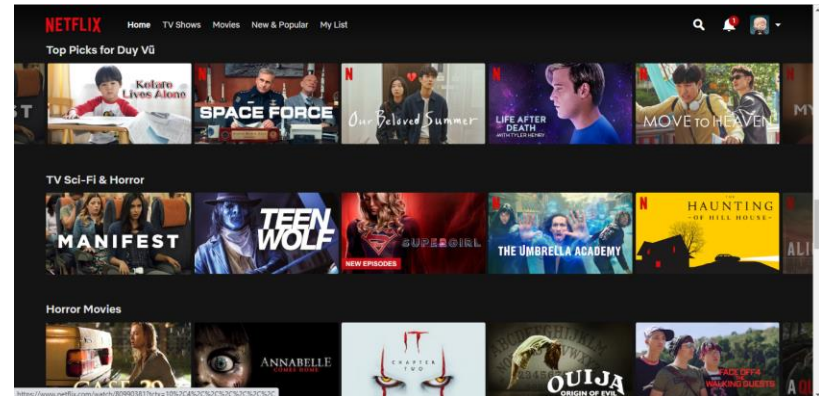
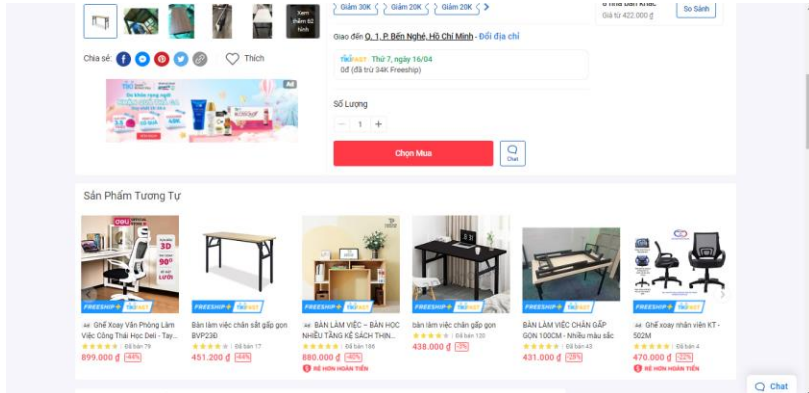
# Phát biểu bài toán

- Đầu vào là đánh giá của người dùng đối với các sản phẩm trước đó trong hệ thống
- Yêu cầu: đưa ra các sản phẩm trong hệ thống mà phù hợp với sở thích của người dùng



# Ý nghĩa của bài toán

- Giúp trải nghiệm của người dùng được cá nhân hóa, do các đề xuất được đưa ra thông qua các đánh giá trước đó của người dùng. Tiết kiệm thời gian cho người dùng trong việc lựa chọn sản phẩm, giúp cho người dùng tập trung hơn vào việc trải nghiệm dịch vụ.
- Đối với các công ty thì sẽ tăng lợi thế cạnh tranh bằng cách nâng cao sự hài lòng của khách hàng. Tăng doanh thu thông qua việc kích thích nhu cầu mua sắm của người dùng bằng các gợi ý sản phẩm.

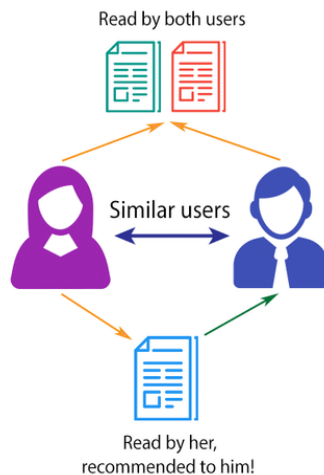


# Các mô hình truyền thống

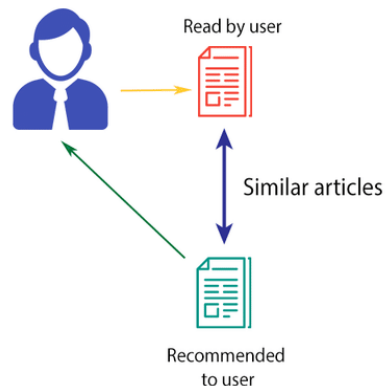
**Phương pháp lọc dựa trên nội dung (Content-base filtering - CB):** Gợi ý sản phẩm cho người dùng các sản phẩm có độ tương đồng cao với các sản phẩm trước đó người dùng đã xem.

**Phương pháp lọc cộng tác (Collaborative Filtering - CF):** Gợi ý sản phẩm cho người dùng dựa trên người dùng giống với họ (được xem xét dựa trên lịch sử xem của người dùng) và gợi ý sản phẩm dựa trên người dùng giống với họ.

COLLABORATIVE FILTERING

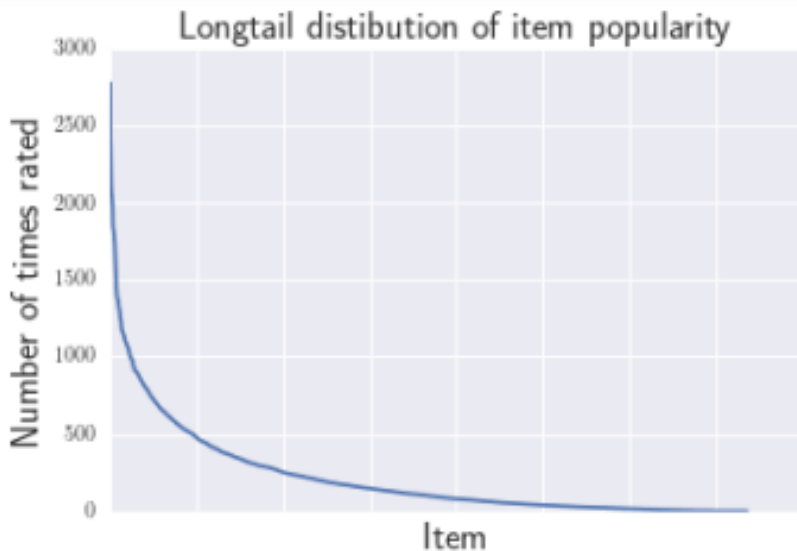


CONTENT-BASED FILTERING



# Điểm hạn chế của mô hình truyền thống

Mặc dù các phương pháp truyền thống cho thấy các kết quả dự đoán tốt trên các tập dữ liệu huấn luyện. Tuy nhiên các dữ liệu được sử dụng để huấn luyện và đánh giá các mô hình đề xuất sản phẩm này lại bị bias. Vì vậy việc huấn luyện trên các tập dữ liệu này sẽ gây ra nhiều vấn đề nghiêm trọng, điển hình như việc khi huấn luyện mô hình cho kết quả tốt nhưng khi đưa vào thực tế thì mô hình lại cho kết quả tệ.



# Điểm hạn chế của mô hình truyền thống

Vì vậy việc ứng dụng các phương pháp suy luận nhân quả vào hệ thống đề xuất sản phẩm sẽ làm giảm đi tác động của dữ liệu bị bias, và có thể đưa ra được các đánh giá chắc chắn rằng liệu hệ thống đề xuất sản phẩm của chúng ta có thật sự cải thiện hay nâng cao trải nghiệm của người dùng hay không.



# Phương pháp suy luận nhân quả

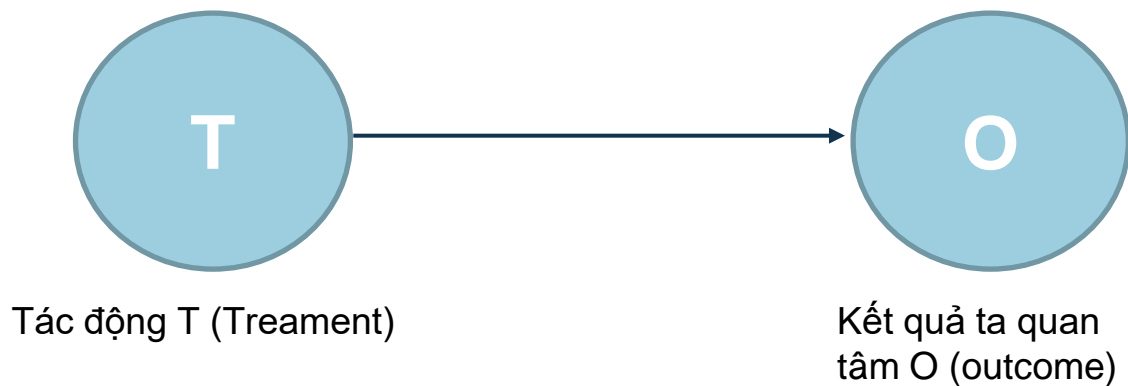
**Phương pháp Chấm điểm theo khuynh hướng nghịch đảo (Inverse Propensity Scoring – IPS):** là phương pháp thường được sử dụng để debias dữ liệu. IPS sẽ đánh lại trọng số của các mẫu dựa trên điểm số xu hướng, điều này làm giảm tác động của các mẫu bị bias. Phương pháp này đề xuất một độ đo chất lượng không bị bias (unbiased quality measument) thay thế cho các độ đo truyền thống để kiểm soát được bias.

$$R(\hat{Y}) = \frac{1}{U \cdot I} \sum_{u=1}^U \sum_{i=1}^I \delta_{u,i}(Y, \hat{Y}).$$

# Các công trình nghiên cứu liên quan

**Phương pháp CauseE:** được tác giả S.Boner và các cộng sự đề xuất trong bài báo “Causal Embeddings for Recommendation”. Nhóm tác giả đã sửa đổi phương pháp phân rã ma trận cổ điển, sử dụng một tập dữ liệu lớn bị bias và một tập dữ liệu nhỏ không bị bias (thu thập được từ chiến lược đề xuất ngẫu nhiên) để tạo đặc trưng cho người dùng và sản phẩm. Thông qua đó cải thiện đáng kể hiệu suất đề xuất trên các cặp người dùng và sản phẩm mới.

# Suy luận nhân quả



# Động lực cho phương pháp nhân quả

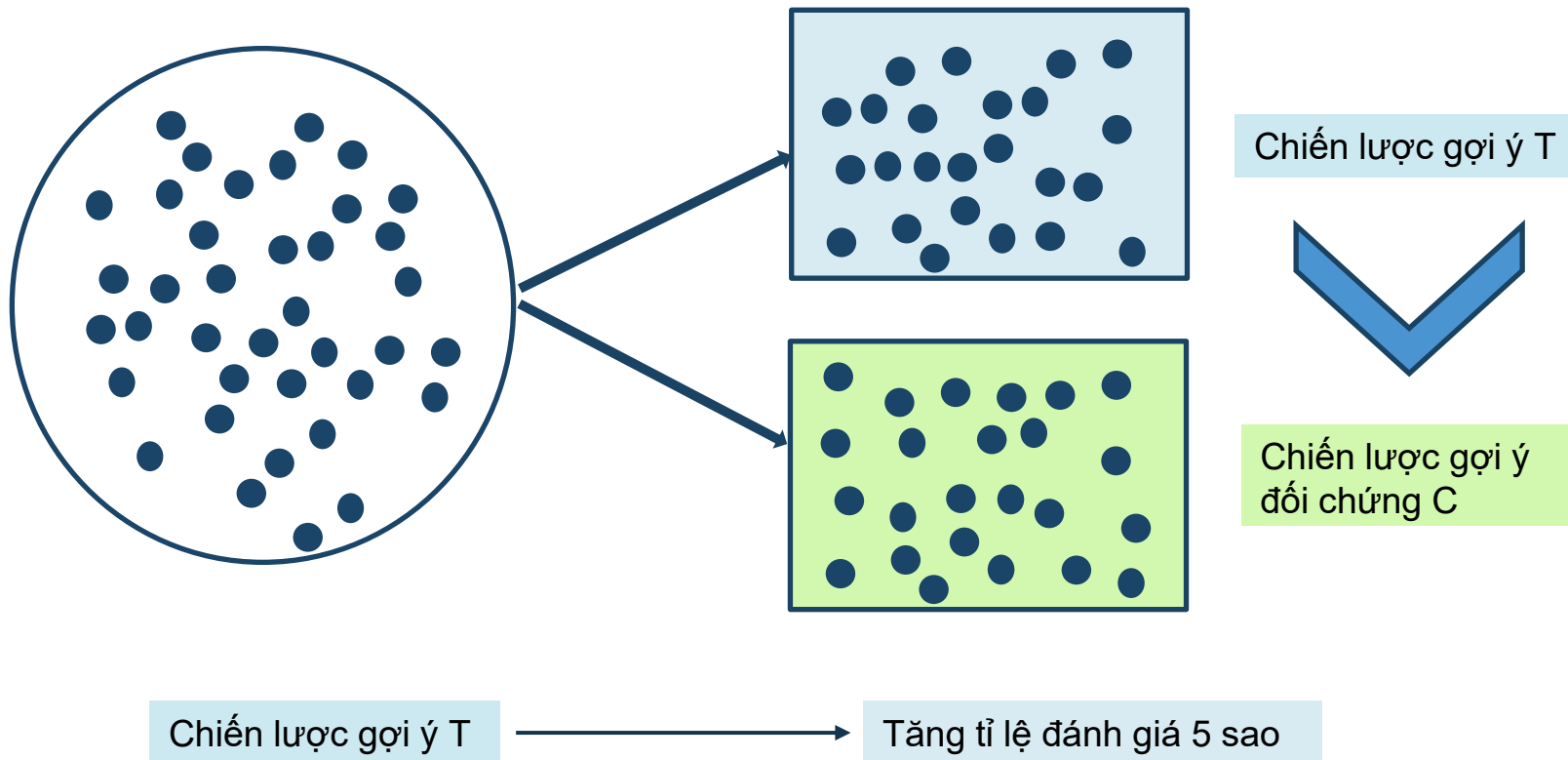


Chiến lược gợi ý T

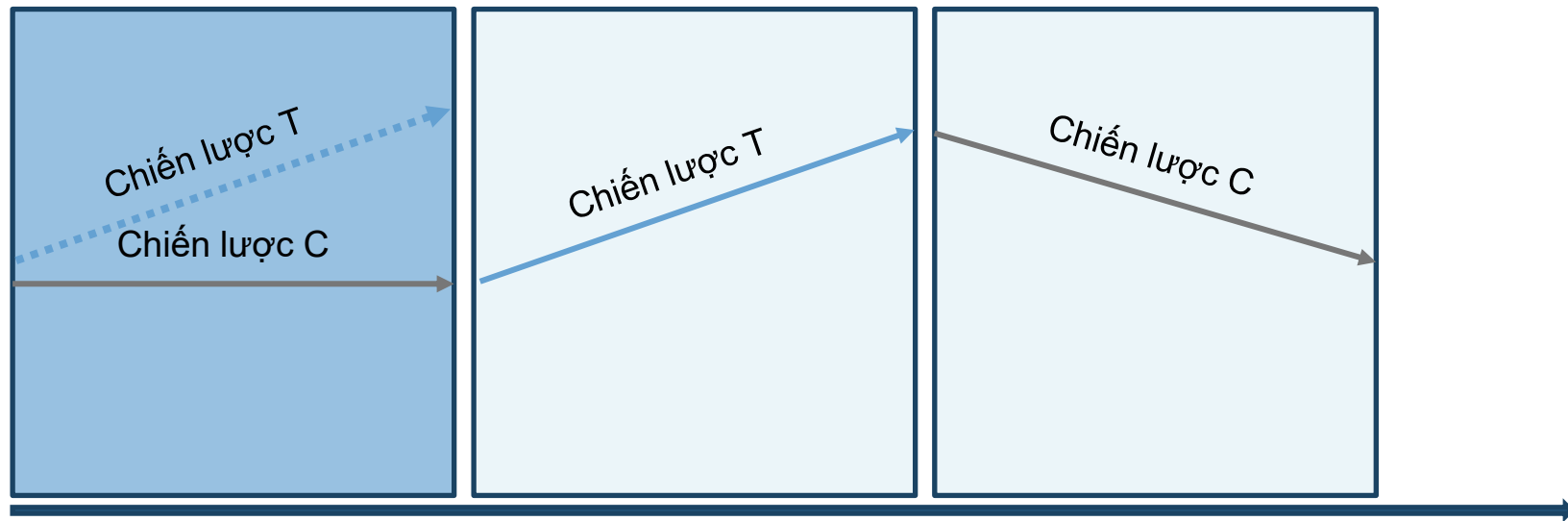
Tăng tỉ lệ đánh giá 5 sao



# Đánh giá hệ thống gợi ý



# Đánh giá hệ thống gợi ý

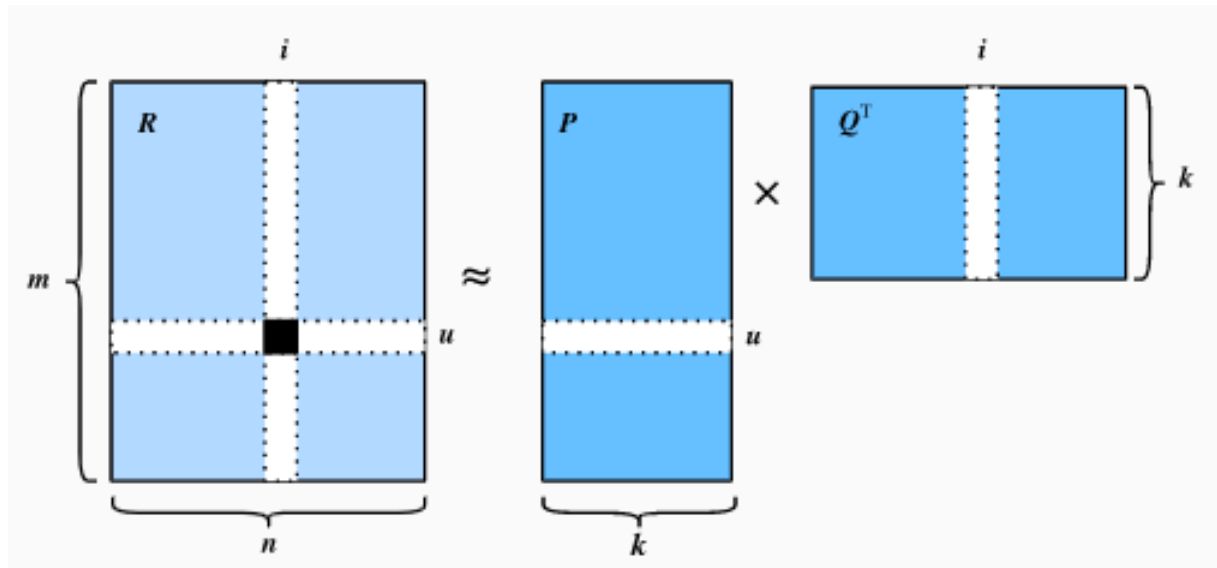


02

# Cơ sở lý thuyết

# Phân rã ma trận (Matrix Factorization)

Phân rã ma trận là một phương pháp thuộc nhóm lọc cộng tác. Mô hình này sẽ phân tích ma trận tương tác của user – item thành tích hai ma trận có hạng thấp hơn, nhằm nắm bắt cấu trúc bậc thấp trong tương tác giữa user – item.





# Phân rã ma trận (Matrix Factorization)

Đánh giá của người dùng  $u$  cho sản phẩm  $i$  được tính theo công thức:

$$\hat{R}_{ui} = p_u q_i^T + b_u + b_i$$

Trong đó:

- $p_u$  là hàng thứ  $u$  trong ma trận đặc trưng người dùng  $P$ .
- $q_i$  là hàng thứ  $i$  trong ma trận đặc trưng của sản phẩm  $Q$ .
- $b_u$  và  $b_i$  là độ lệch của người dùng và sản phẩm.

Sau đó mô hình sẽ được huấn luyện bằng cách cực tiểu hóa trung bình bình phương sai số giữa đánh giá dự đoán và đánh giá thực. Hàm mục tiêu được định nghĩa như sau:

$$\operatorname{argmin}_{\mathbf{P}, \mathbf{Q}, b} \sum_{(u,i) \in \mathcal{K}} \|\mathbf{R}_{ui} - \hat{\mathbf{R}}_{ui}\|^2 + \lambda(\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2 + b_u^2 + b_i^2)$$

# Hiệu quả tác động lên cá thể

- Một trong các phương pháp đánh giá ảnh hưởng của yếu tố tác động lên kết quả ta quan tâm là đo lường hiệu quả tác động lên cá thể (Individual Treatment Effect - ITE), được tính theo công thức như sau:  $ITE = Y_{1i} - Y_{0i}$

Trong đó:

- $Y_{0i}$  là kết quả tiềm năng của cá thể  $i$  không được can thiệp (untreatment/control)
- $Y_{1i}$  là kết quả tiềm năng trên cùng cá thể  $i$  khi tiếp nhận can thiệp

Ví dụ: trong quá trình giãn cách xã hội, ta muốn xem hình thức học trực tuyến có ảnh hưởng như thế nào đến một sinh viên thì  $ITE = \text{điểm trung bình kết quả học trực tuyến} - \text{điểm trung bình kết quả học tập học trực tiếp của sinh viên đó}$ .

# Chiến lược đề xuất

Giả định  $\Pi_x$  là chiến lược đề xuất ngẫu nhiên (stochastic policy) nó liên kết mỗi người dùng  $u_i$  và sản phẩm  $p_j$  một xác suất cho người dùng  $u_i$  được hiển thị đề xuất sản phẩm  $p_i$ .

$$p_j \sim \pi_x(.|u_i)$$

# Phần thưởng cho chiến lược đề xuất

$r_{ij}$  là phần thưởng/ đầu ra thực trong việc đề xuất sản phẩm  $p_j$  cho người dùng  $u_i$ :

$$r_{ij} \sim r(.|u_i, p_j)$$

$y_{ij}$  là phần thưởng/ đầu ra quan sát được cho mỗi cặp người dùng  $u_i$  và sản phẩm  $p_j$  với chiến lược gợi ý  $\Pi_i$

$$y_{ij} = r_{ij}\pi_x(p_j|u_i)$$

Phần thưởng  $R^{\pi_x}$  của chiến lược gợi ý  $\pi_x$  được tính bằng tổng phần thưởng của tất cả người dùng

$$R^{\pi_x} = \sum_{ij} r_{ij}\pi_x(p_j|u_i)p(u_i) = \sum_{ij} y_{ij}p(u_i) = \sum_i R_{ij}$$

# Individual Treatment Effect (ITE)

Giá trị ITE của một chiến lược gợi ý  $\pi_x$  với mỗi người dùng  $u_i$  và sản phẩm  $p_j$  cho trước, được tính bằng sự khác biệt giữa phần thưởng của chiến lược  $\pi_x$  và phần thưởng của chiến lược gợi ý đối chứng  $\pi_c$

$$ITE_{ij}^{\pi_x} = R_{ij}^{\pi_x} - R_{ij}^{\pi_c}$$

Nhiệm vụ của ta là tìm một chiến lược gợi ý  $\Pi^*$  sao cho tổng của ITE lớn nhất

$$\pi^* = \arg \max_{\pi_x} \{ITE^{\pi_x}\}$$

Trong đó

$$ITE^{\pi_x} = \sum_{ij} ITE_{ij}^{\pi_x}$$

# Chiến lược đề tối ưu ITE

Đối với bất kì chiến lược gợi ý đối chứng  $\pi_c$  nào, chiến lược cải tiến tốt nhất  $\Pi^*$  là chiến lược hiển thị đến mỗi người dùng sản phẩm mà có phần thưởng tương ứng cao nhất

$$\pi_{det} = \begin{cases} 1, & \text{if } p_j = p_i^* \\ 0, & \text{otherwise} \end{cases}$$

Trong đó:  $\Pi^* = \Pi_{det}$

$$p_i^* = \arg \max_{p_j} r_{ij}$$

Và  $\pi_{det}$  là một chiến lược gợi ý xác định (deterministic policy) để hiển thị sản phẩm phù hợp nhất cho mỗi người dùng

# Các phương pháp tìm $\Pi^*$

Để tìm chiến lược tối ưu  $\Pi^*$ , chúng ta cần tìm những sản phẩm có phần thưởng được cá nhân hóa  $r_i^*$  cao nhất cho mỗi người dùng  $u_i$

Trong thực tế, ta không thể quan sát trực tiếp được phần thưởng  $r_{ij}$ , tuy nhiên như đã đề cập ở phần trên:

$$y_{ij} = r_{ij} \pi_x(p_j | u_i)$$

Phương pháp dựa trên IPS (Inverse Propensity Scoring) để dự đoán phần thưởng  $r_{ij}$  không quan sát được:

$$\hat{r}_{ij} \approx \frac{y_{ij}}{\pi_c(p_j | u_i)}$$

Tuy nhiên các phương pháp này lại gặp vấn đề trong việc đối với các sản phẩm có xác suất hiển thị thấp lại dẫn đến phần thưởng cao



03

# Phương pháp



# Phương pháp tìm hiểu

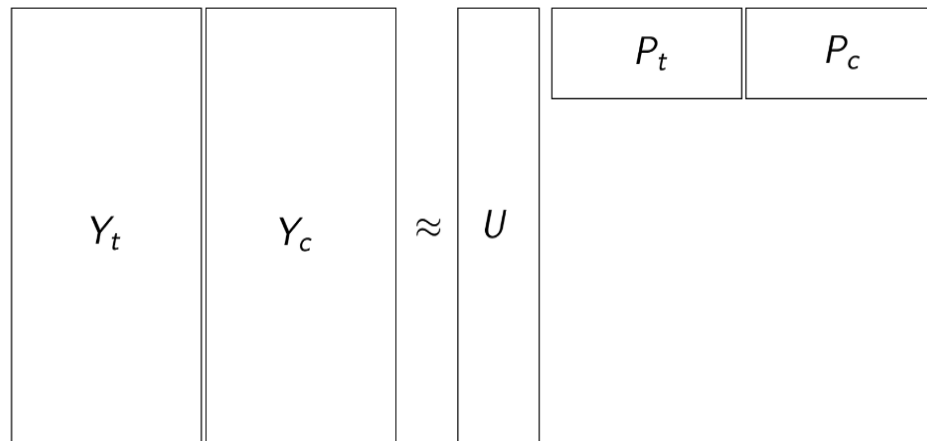
- Từ vấn đề về bias của tập dữ liệu, tác giả tập trung vào việc xây dựng một hệ thống gợi ý giúp người dùng có thể tiếp cận ngẫu nhiên với một sản phẩm.
- Tuy nhiên một hệ thống gợi ý ngẫu nhiên thì rõ ràng không đủ tốt trong việc thay đổi hành vi người dùng, hơn nữa mô hình sẽ không ổn định nếu chỉ có một lượng dữ liệu nhỏ.
- Do vậy, trong framework của mô hình nhân quả này, tác giả xây dựng một công cụ dự đoán dựa trên giả định rằng tồn tại hai tập dữ liệu huấn luyện

# Phương pháp tìm hiểu

- Một tập dữ liệu nhỏ  $S_T$  thu thập các tương tác của người dùng với sản phẩm được tạo ra bằng một chiến lược gợi ý ngẫu nhiên  $T$  (chiến lược này gợi ý bất kì bộ phim nào đến mọi người dùng với xác suất như nhau)
- Một tập dữ liệu rất lớn  $S_C$  thu thập tương tác của người dùng với sản phẩm được tạo ra bằng một chiến lược gợi ý đối chứng  $C$  (control recommendation policy)

# Xây dựng chiến lược gợi ý đa tác vụ

Tác giả đề xuất mục tiêu đa tác vụ để kết hợp việc phân rã ma trận của các quan sát  $y_{ij}^c$  thuộc  $S_c$  và ma trận của các quan sát  $y_{ij}^t$  thuộc  $S_t$



# Xây dựng chiến lược gợi ý đa tác vụ cho tập người dùng cố định

- Để đơn giản, ta giả định rằng các chiến lược gợi ý ngẫu nhiên  $T$  và chiến lược gợi ý đối chứng  $C$  được sử dụng trên cùng một tập người dùng
- Sự khác biệt của hai chiến lược được giải thích dựa trên sự khác nhau giữa các sản phẩm được gợi ý
- Tác giả giả định một chiến lược gợi ý độc lập với người dùng và đưa ra một mục tiêu đa tác vụ để giải thích việc kết hợp giữa tập dữ liệu nhỏ  $S_t$  và tập dữ liệu lớn  $S_c$  sau đó tính toán sự khác biệt giữa hai kết quả đại diện của các sản phẩm

# Dự đoán phần thưởng

- Để dễ dàng thực hiện mục tiêu này, tác giả giả định rằng phần thưởng của chiến lược gợi ý T và chiến lược gợi ý đối chứng C đều có thể được ước lượng bằng một hàm tuyến tính dựa trên các đại diện người dùng cố định  $u_i$ :

$$y_{ij}^c \approx \langle \theta_j^c, u_i \rangle$$

$$y_{ij}^t \approx \langle \theta_j^t, u_i \rangle$$

- Với  $\Theta_j$  là vector đại diện cho bộ sản phẩm j, và  $\langle x, y \rangle$  thể hiện cho phép tích vô hướng của 2 vector x, y
- Từ đó, ta có thể ước lượng được ITE của từng cặp người dùng - sản phẩm i, j

$$\widehat{ITE}_{ij} = \langle \theta_j^t, u_i \rangle - \langle \theta_j^c, u_i \rangle = \langle \theta_j^\Delta, u_i \rangle$$

# Xây dựng hàm lỗi

Phần đầu tiên của mục tiêu dự đoán kết hợp là dự đoán có giám sát cho  $y_{ij}^t$  được huấn luyện trên tập  $S_t$  được hiểu diễn bởi  $l_{ij}^t$  theo công thức sau:

$$l_{ij}^t = L(\langle \theta_j^t, u_i \rangle, y_{ij}^t) + \Omega(\theta_j^t)$$

Trong đó:  $L$  là một hàm lỗi bất kỳ và  $\Omega(\cdot)$  là regularizator  
Chuyển sang ký hiệu ma trận ta có công thức sau:

$$L_t = \sum_{(i,j,y_{ij}) \in S_t} l_{ij}^t = L(U\Theta_t, Y_t) + \Omega(\Theta_t)$$

Trong đó:

- $U$  là ma trận biểu diễn cho tập người dùng cố định
- $\Theta_t$  là ma trận tham số biểu diễn cho các sản phẩm theo chiến lược gợi ý  $T$

# Xây dựng hàm lỗi

Tương tự,  $L_c$  được tính theo công thức:

$$L_c = L(U\Theta_c, Y_c) + \Omega(\Theta_c)$$

Để tận dụng phần  $\theta_j^t$  đã tính toán trước đó, ta có thể chuyển các biểu diễn cho các sản phẩm đối chứng theo  $\theta_i^t$  và  $\theta_i^\Delta$  theo công thức:

$$\langle \theta_j^c, u_i \rangle = \langle \theta_j^t, u_i \rangle - \langle \theta_j^\Delta, u_i \rangle$$

Từ đó,  $l_{ij}^c$  sẽ được biểu diễn như sau:

$$l_{ij}^c = L(\langle \theta_j^t - \theta_j^\Delta, u_i \rangle, y_{ij}^c) + \Omega(\theta_j^t) + \Omega(\theta_j^\Delta)$$

# Kết hợp hai tác vụ

Sau đó ta kết hợp hai tác vụ tính toán  $L_c$  và  $L_t$  để tạo ra công thức tính độ lỗi cho  $L_{CausE}^{prod}$  như sau:

$$L_{CausE}^{prod} = \underbrace{L(U\Theta_t, Y_t) + \Omega(\Theta_t)}_{\text{treatment task loss}} + \underbrace{L(U\Theta_c, Y_c) + \Omega(\Theta_c)}_{\text{control task loss}} + \underbrace{\Omega(\Theta_t - \Theta_c)}_{\text{regularizer between tasks}}$$

Ta định nghĩa tương tự một hàm lỗi  $L_{CausE}^{user}$  với người dùng được học qua tập các sản phẩm cố định. Loại mô hình hóa này phù hợp với những trường hợp mà chiến lược tác động và chiến lược kiểm soát khác nhau về sự lựa chọn của người dùng



# Xây dựng chiến lược gợi ý đa tác vụ cho tập người dùng không cố định

Trong trường hợp này, sự thay đổi về việc tiếp cận đến những bộ phim giữa chiến lược tác động và chiến lược đối chứng được giải thích bằng cả sự lựa chọn khác nhau ở những người dùng và các sản phẩm được hiển thị. Để thực hiện điều này, tác giả đã điều chỉnh hàm mục tiêu ở phần trước, cho phép tập người dùng có thể thay đổi, hàm lỗi cho phương pháp này được tính bằng  $L_{\text{CausE}}$

$$\begin{aligned} L_{\text{CausE}} = & \underbrace{L(\Gamma_t \Theta_t, Y_t) + \Omega(\Gamma_t, \Theta_t)}_{\text{treatment task loss}} + \underbrace{L(\Gamma_c \Theta_c, Y_c) + \Omega(\Gamma_c, \Theta_c)}_{\text{control task loss}} + \\ & + \underbrace{\Omega(\Gamma_t - \Gamma_c) + \Omega(\Theta_t - \Theta_c)}_{\text{regularizers between tasks}} \end{aligned}$$

04

# Kết quả thí nghiệm