



Bài giảng môn học:

Kỹ nghệ tri thức và học máy (4080540)

CHƯƠNG 5: HỆ THỐNG GỢI Ý (Recommender System)

Giảng viên: Đặng Văn Nam

Email: dangvannam@humg.edu.vn

Nội dung chương 5



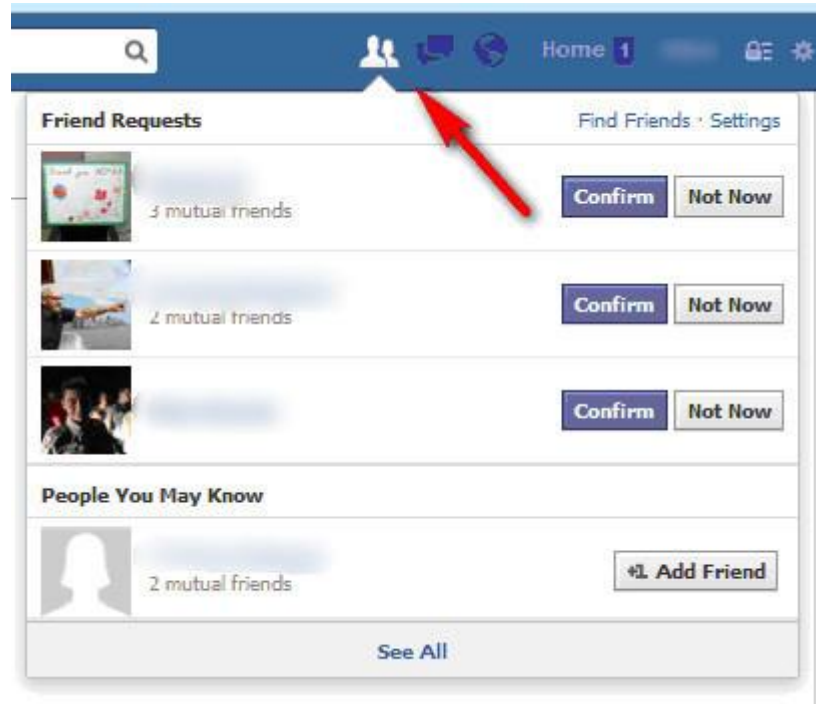
- 1. Giới thiệu chung**
- 2. Phân loại hệ thống đề xuất**
- 3. Các phương pháp tính toán độ tương đồng**
- 4. Thách thức**
- 5. Ví dụ minh họa**

1. Giới thiệu

Giới thiệu



- Chúng ta sử dụng Facebook và mới chỉ kết bạn với một vài người. Tuy nhiên vài hôm sau, **Facebook đã tự gợi ý** cho chúng ta những người bạn khác nhau mà thậm chí ngay cả chúng ta cũng không biết họ???



- Amazon Recommendation Services
○○●○○○○○○○○○○○○○○○○○○

123Mua.vn Recommendation Services
○○○○○○○○○○○○○○○○○○○○


Business Model

Recommendation Services (cont.)

Purchasing Product Recommendations:

Shopping Cart

Items to buy now



Audio Video RCA Cable 12ft. cable

In Stock

Shipped from: Annamex

Gift options not available (Learn more)

Delete • Save for later

Price

Quantity

\$3.62

1

Recommendations based on Shopping Cart


Subtotal: \$3.62

The price and availability of items at Amazon.com are subject to change. The shopping cart is a temporary place to store a list of your items and reflects each item's most recent price. [Learn more](#)

Do you have a gift card or promotional code? We'll ask you to enter your claim code when it's time to pay.


Customers Who Bought Items in Your Shopping Cart Also Bought

Page 1 of 5




Cable 1

40650 RCA...




Cable 2

40650 RCA...



Cable 3


VGA Adapter to TV S-Video RCA...



Cable 4

eForCity PC Computer VGA to TV...

Customers Who Bought Audio Video RCA Cable 12ft. Also Bought




Female to Female 3-RCA Coupler

★★★★★ (50)

\$1.95

Add to Cart




3 RCA PHONO FEMALE SOCKET...

★★★★★ (18)

\$0.75


Add to Cart



eForCity VGA to TV S-Video RCA...

\$2.13

Add to Cart



S-Video Cable 12 Ft. Gold Plated

★★★★★ (53)

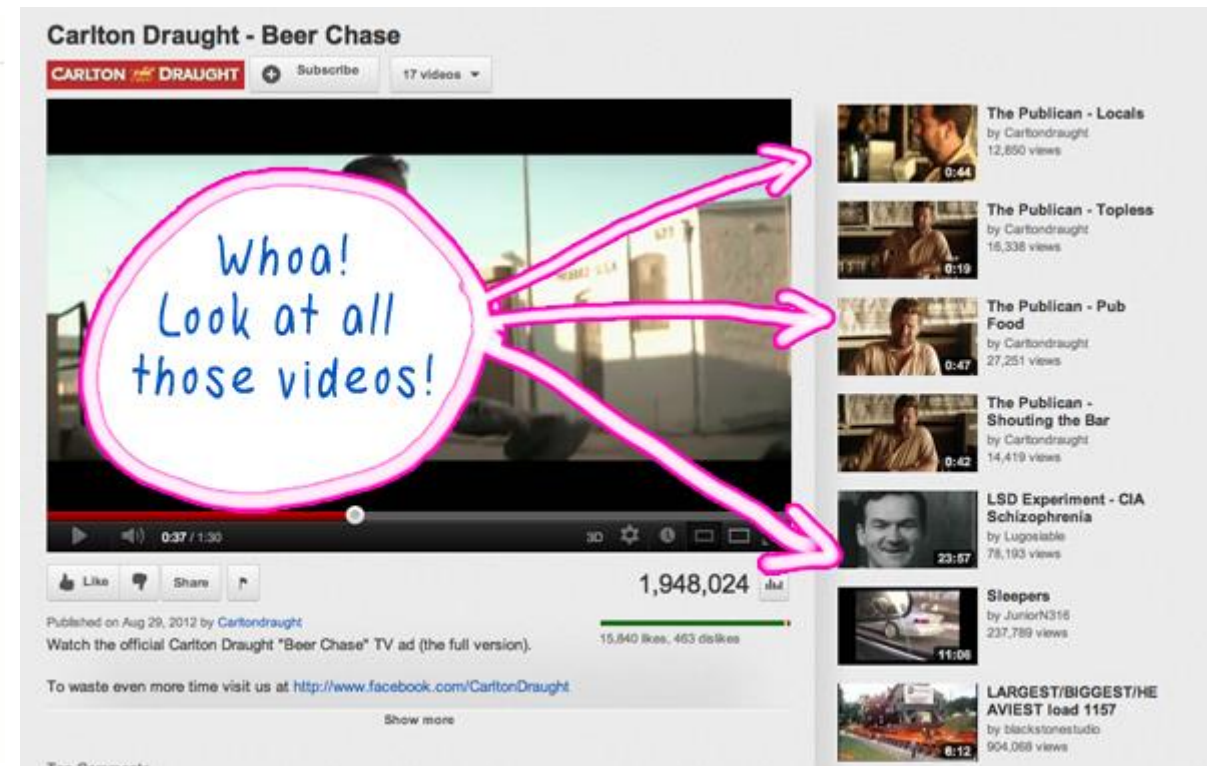
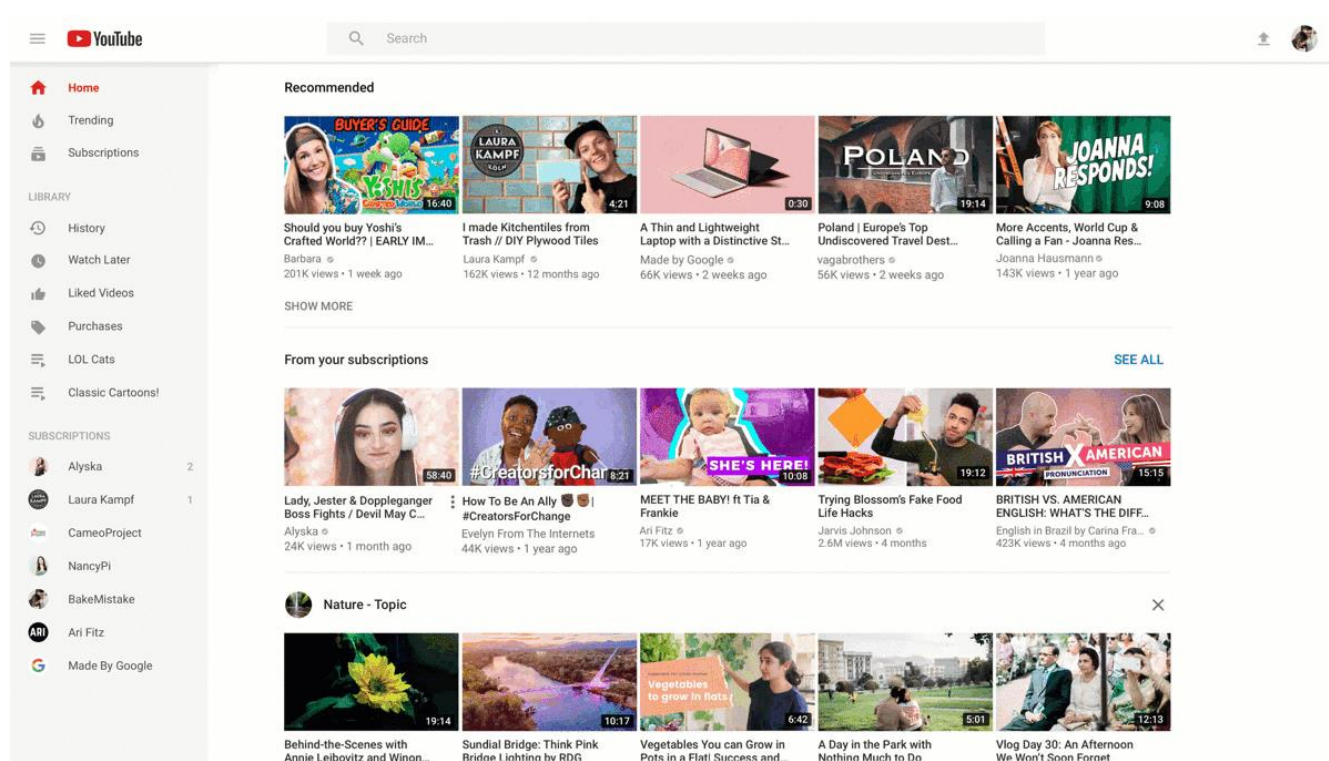
\$2.85

Add to Cart

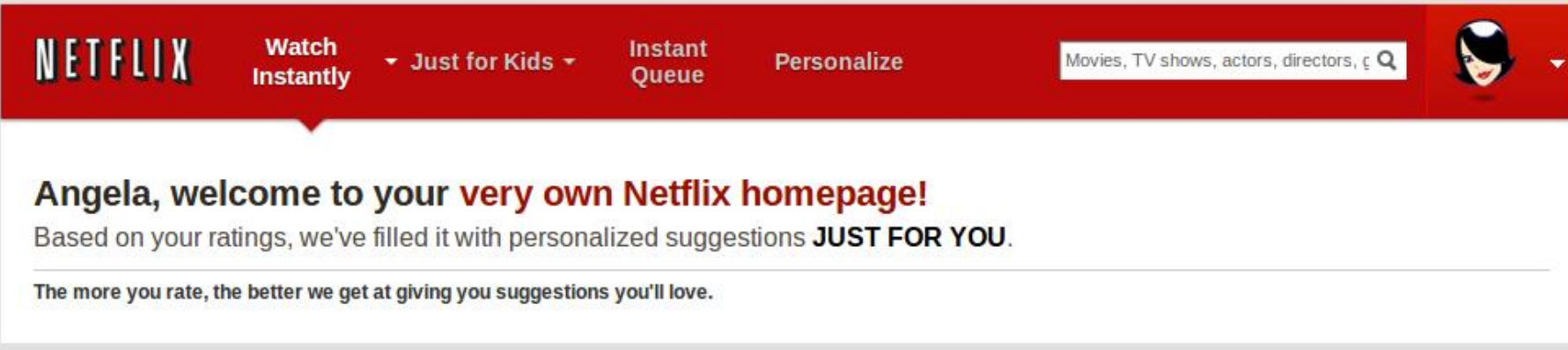
Giới thiệu



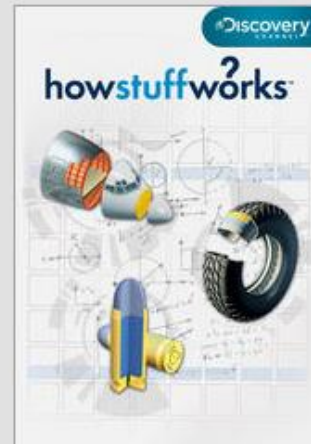
- Chúng ta mở **Youtube** và tìm kiếm một bản nhạc vàng, nằm xuống giường và thưởng thức âm nhạc, những bài hát tiếp theo sẽ được Youtube thân yêu tự động gợi ý và tất nhiên rất hiếm khi nó gợi ý một bản nhạc rock cho bạn trong khi bạn đang nghe nhạc vàng phải không nào ???



Giới thiệu



Top 10 for Angela



Movies Popular on Netflix



Everything is a Recommendation



Over 80% of what people watch comes from our recommendations

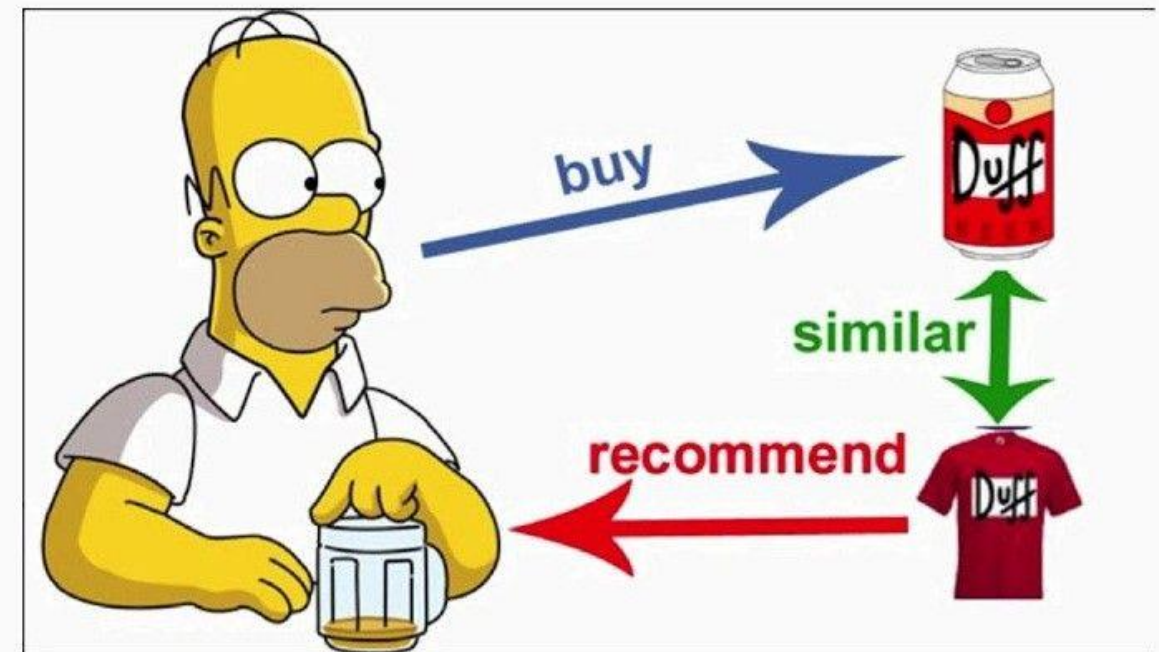
Recommendations are driven by **Machine Learning**

Giới thiệu

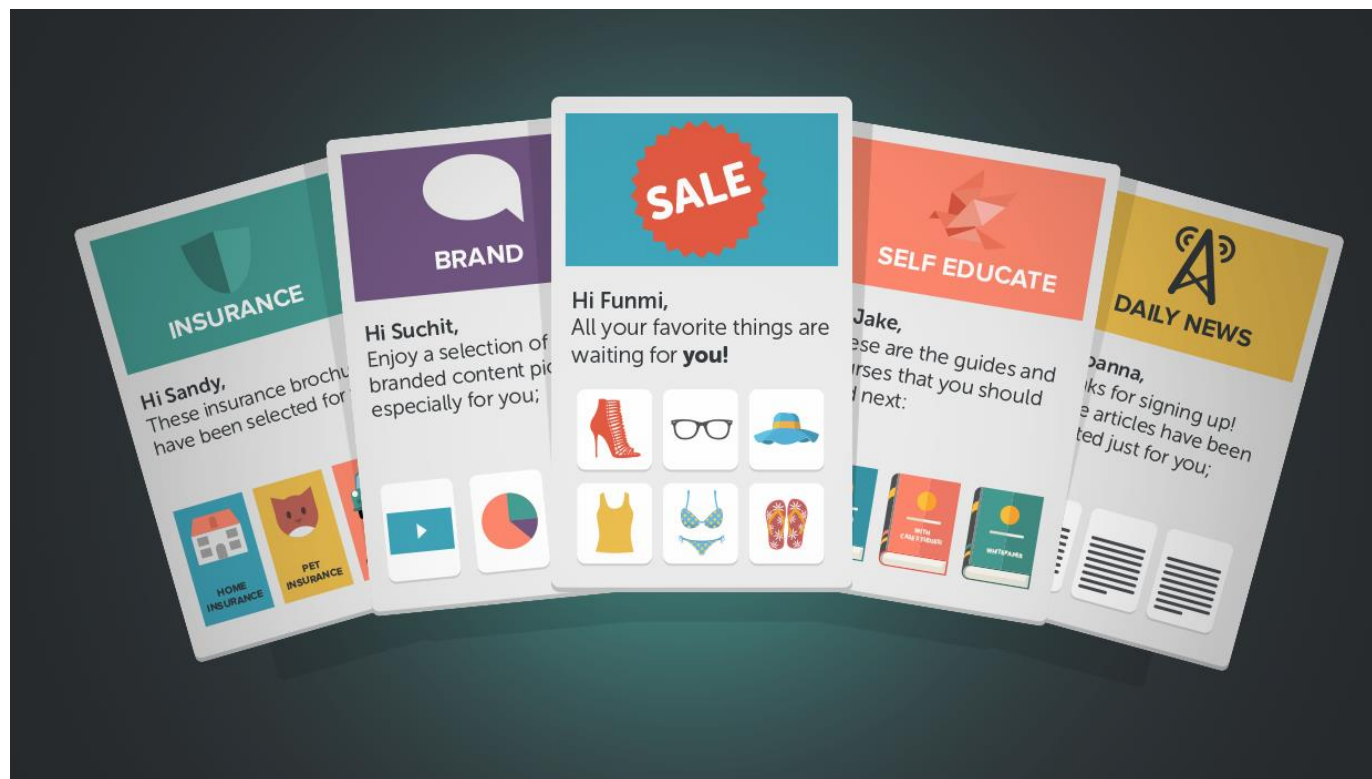


Hệ thống gợi ý (Recommender Systems)

- Hệ thống gợi ý (Recommender Systems) là một thành phần trong hệ thống thông tin. Mục đích của nó là hỗ trợ người dùng tìm kiếm được đúng thông tin cần thiết, dự đoán sở thích hay xếp hạng mà người dùng có thể dành cho một mục thông tin (item) nào đó mà họ chưa xem xét tới trong quá khứ.
- Các gợi ý được đưa ra là kết quả của việc tính toán dựa trên việc thu thập dữ liệu về người dùng như thói quen, sở thích, hành vi khi mua hàng, khi đưa ra các đánh giá cá nhân...
- Việc thực hiện tính toán được xây dựng trên các thuật toán Học máy, đưa ra các dự đoán tốt nhất về sản phẩm mà người dùng có thể thích, giúp gia tăng số lượng sản phẩm bán được.



Hệ thống gợi ý (Recommender Systems)



- Hệ thống gợi ý đang là một công cụ mạnh mẽ, được ứng dụng chủ yếu cho các mạng xã hội, giải trí trực tuyến, thương mại điện tử....

Lợi ích mang lại là gì?

Hệ thống gợi ý (Recommender Systems)

- Tăng doanh thu, tăng CRT (Click-through rate)...
- Tăng mức tín nhiệm và trung thành của khách hàng.
- Thêm hiểu biết về khách hàng.
- Khả năng đưa ra các dịch vụ các nhân hóa, hướng tới từng đối tượng khách hàng cụ thể.

Why Netflix thinks its personalized recommendation engine is worth \$1 billion per year



Nathan McAlone  
🕒 14.06.2016, 21:36 💬 228



FACEBOOK



LINKEDIN



TWITTER



EMAIL



PRINT

After a long refinement process, Netflix finally released its first “global” recommendation engine in December.

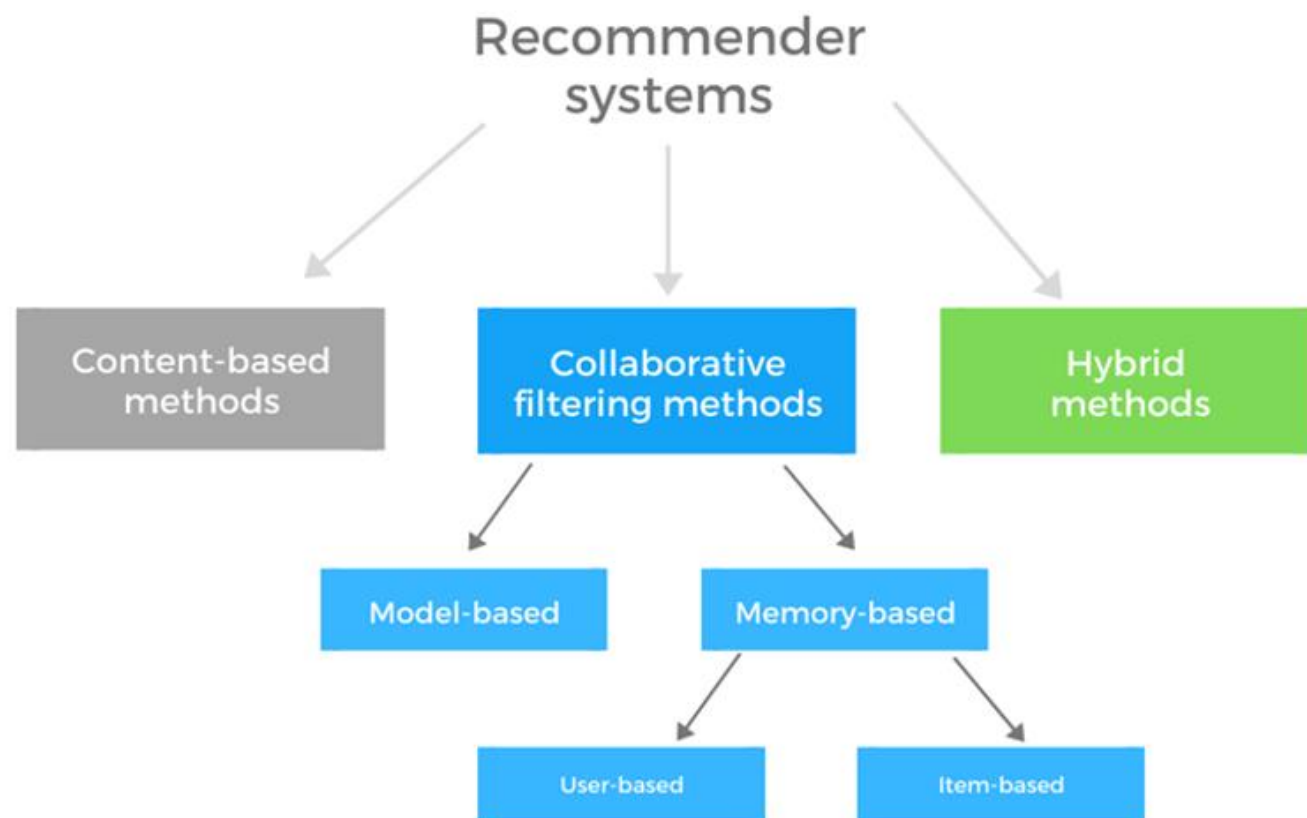
The engine takes dozens of algorithms into account and compares you with similar users in the more than 190 countries where Netflix's service is available.



2. Phân loại hệ thống gợi ý

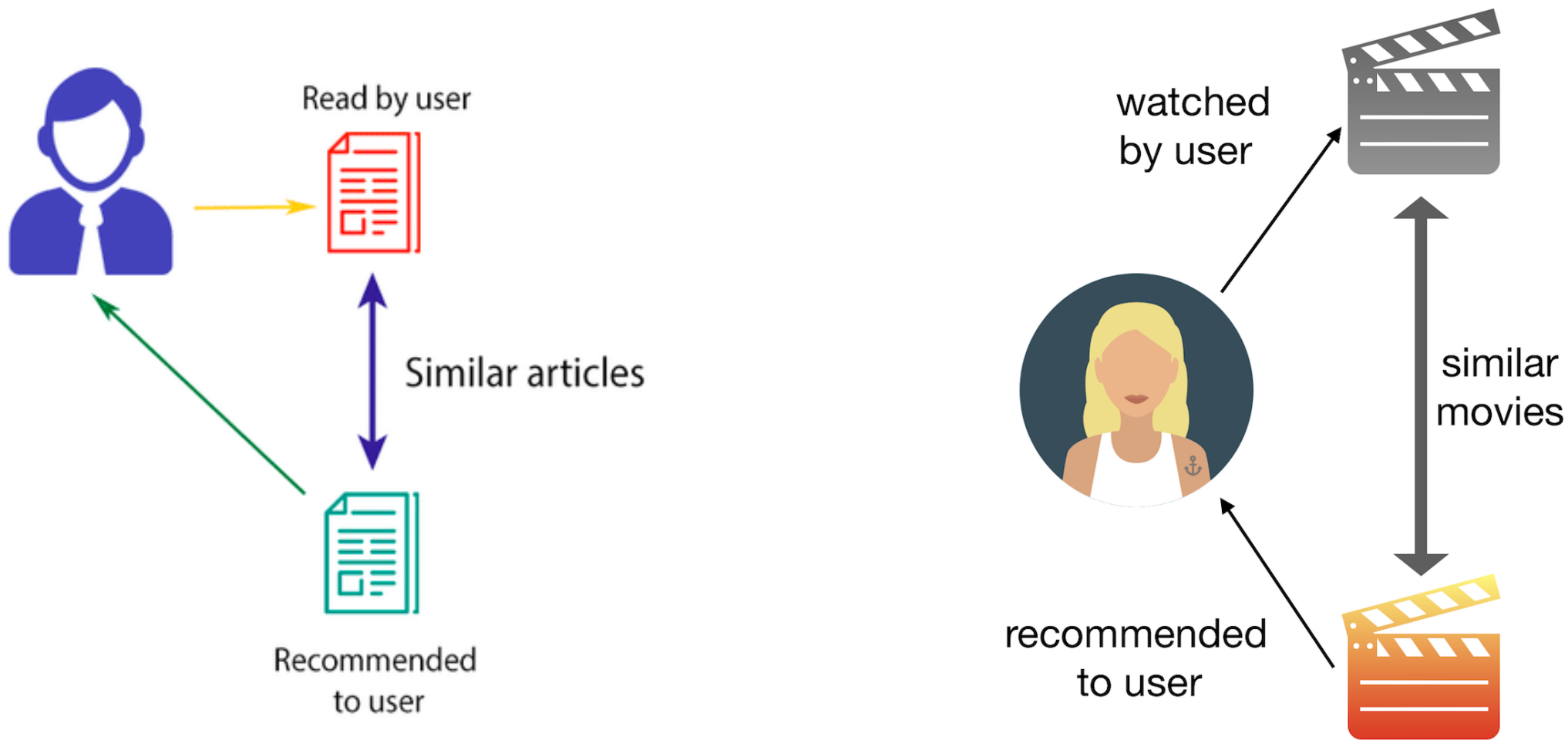
Phân loại các hệ thống gợi ý

1. Hệ thống gợi ý dựa trên nội dung (**Content based recommender systems**)
2. Hệ thống gợi ý dựa trên các user – Lọc cộng tác (**Collaborative filtering recommender systems**)
3. Hệ thống gợi ý lai (**Hybrid systems**)



Hệ thống gợi ý dựa trên nội dung

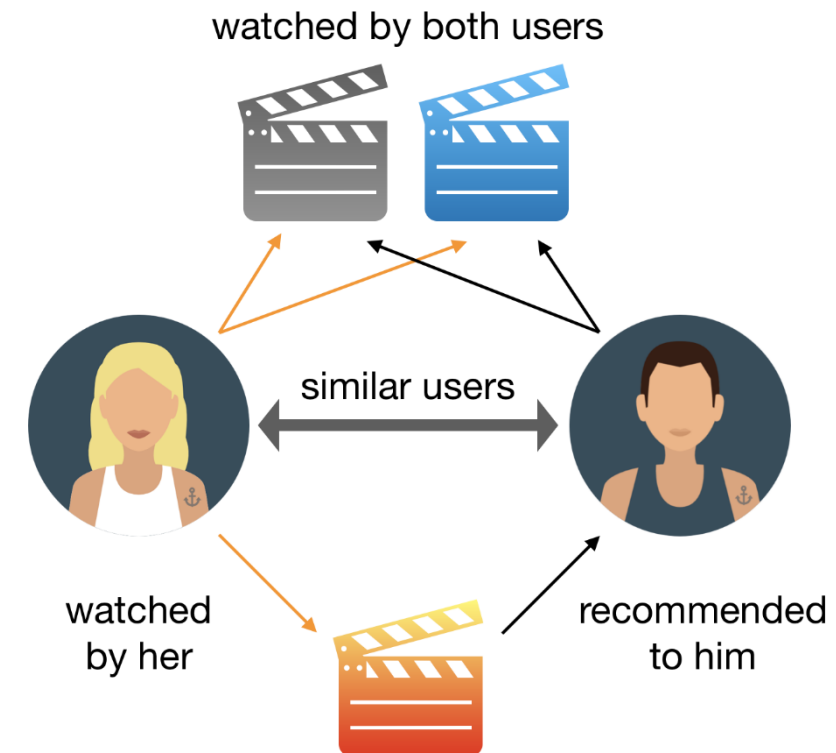
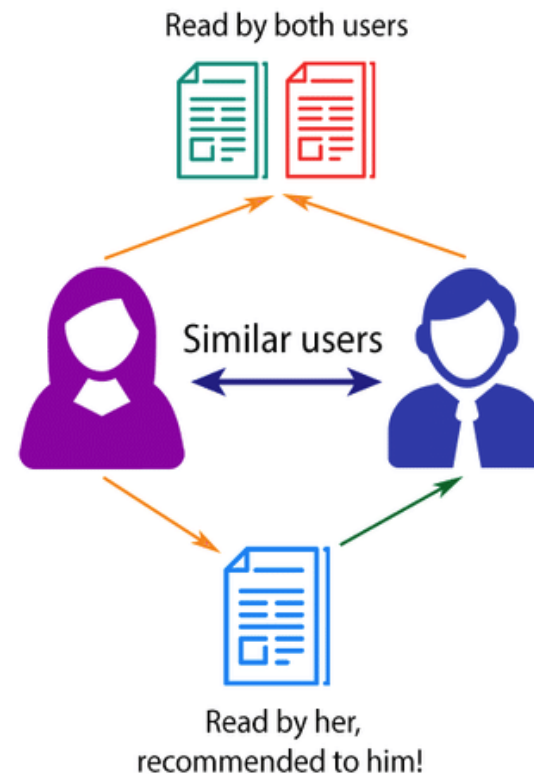
Hệ thống dựa trên nội dung, đặc điểm của mục tin, sản phẩm, video...hiện tại và sau đó gợi ý cho người dùng các mục tin, sản phẩm, video... tương tự.



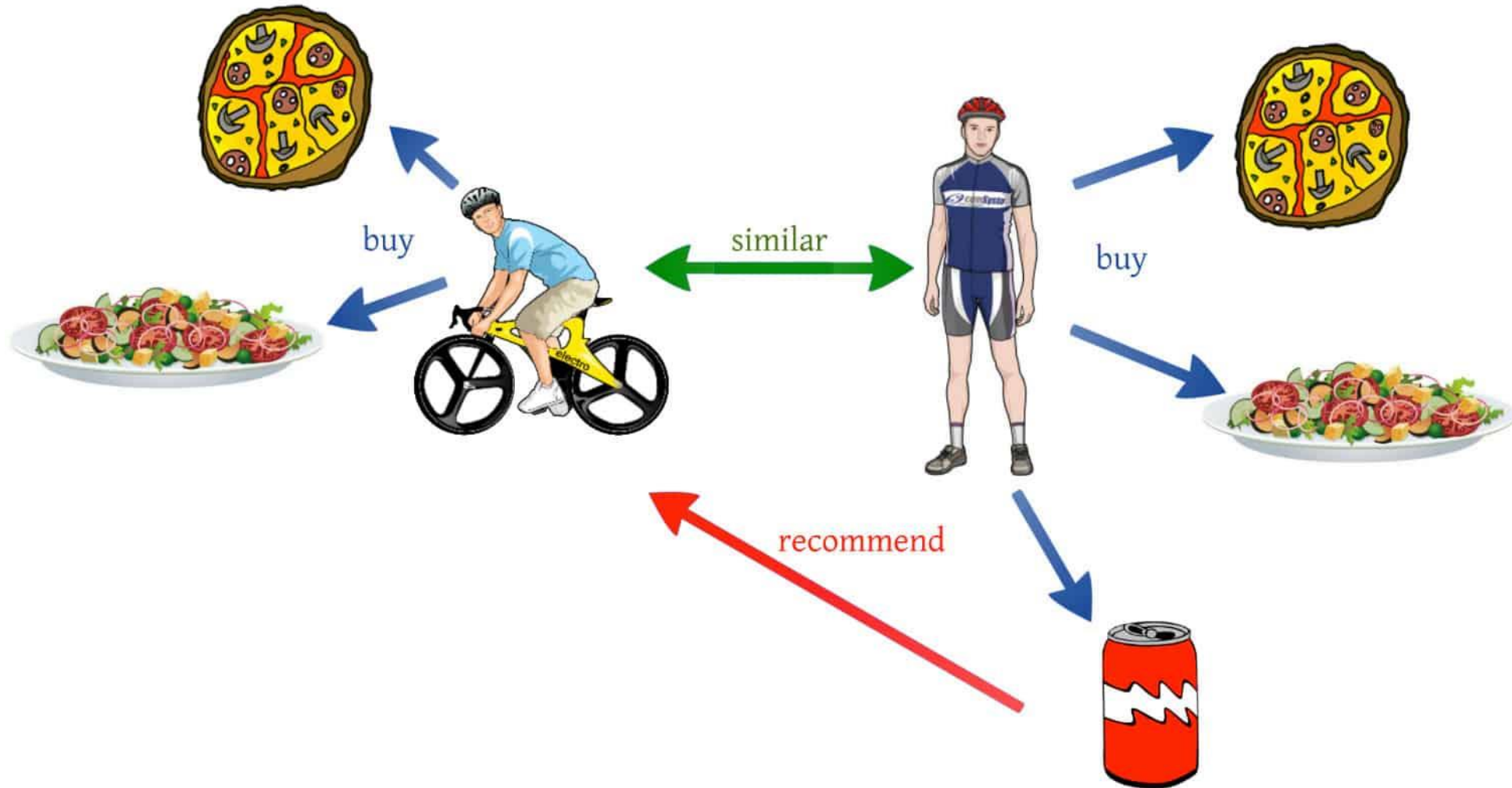
Hệ thống gợi ý dựa trên user – lọc cộng tác

Hệ thống sẽ phân tích các user có cùng đánh giá, cùng mua **mục tin** hiện tại. Sau đó tìm ra danh sách các **mục tin** khác cũng được đánh giá bởi các user này, xếp hạng và gợi ý cho người dùng. Tư tưởng của phương pháp này chính là dựa trên **sự tương đồng về sở thích giữa các người dùng** để đưa ra các gợi ý.

Về bản chất, Nó lọc trên những người có cùng sở thích, hay những người có cùng những hành vi tương tự, cùng bấm like, cho điểm đối với cùng một item.



Hệ thống gợi ý dựa trên user – lọc cộng tác

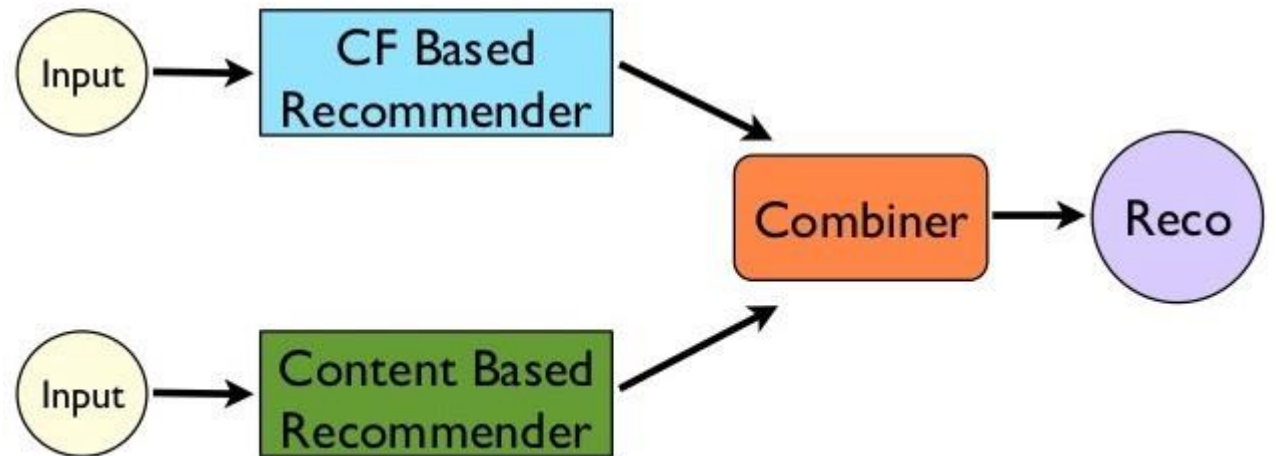


Hệ thống gợi ý lai



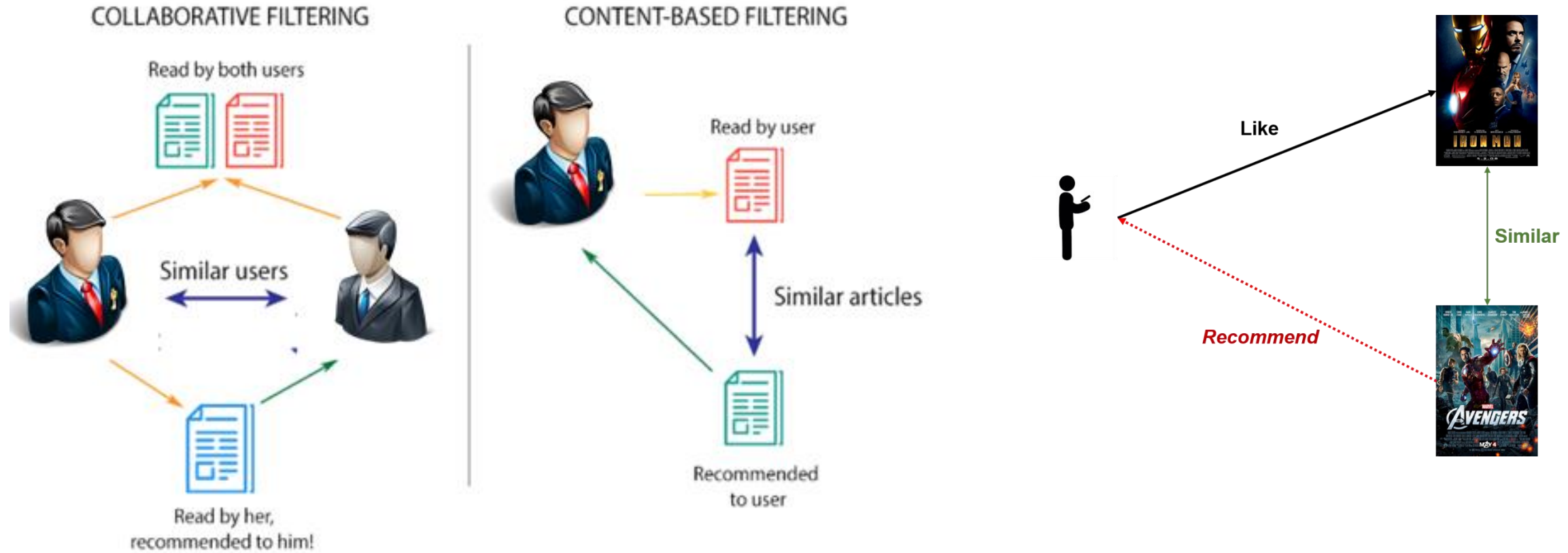
Kết hợp các ý tưởng của Content-based recommender và Collaborative filtering để xây dựng một hệ thống gợi ý.

Hybrid Recommendations



3. Các phương pháp đánh giá độ tương đồng

Các phương pháp tính toán độ tương tự



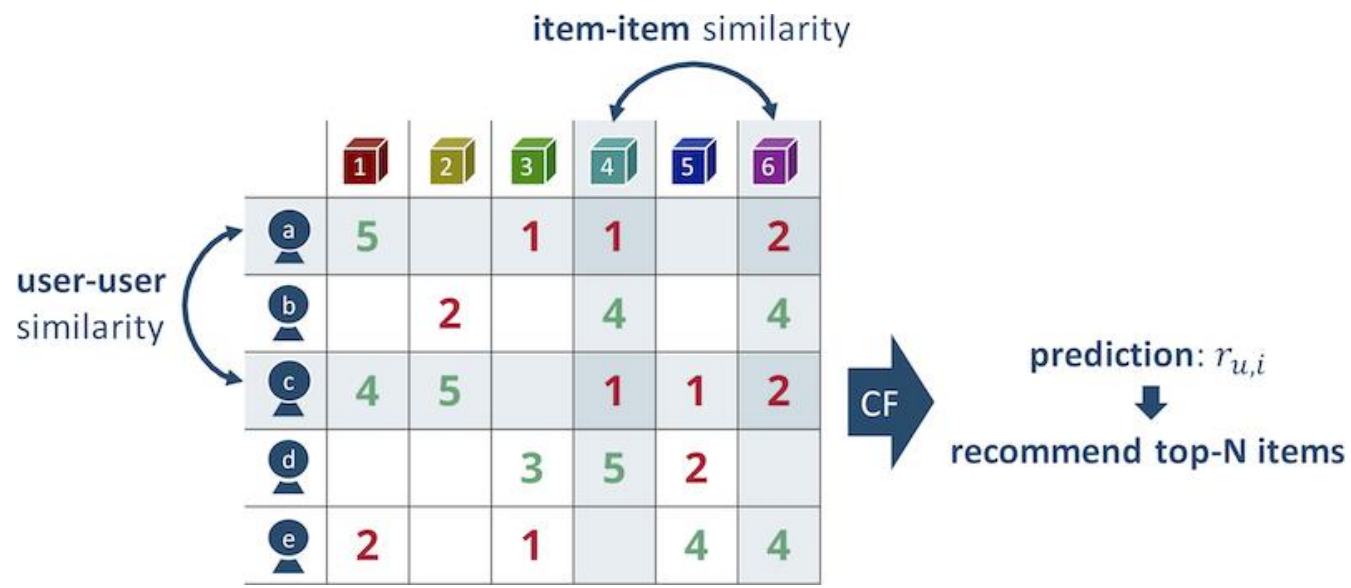
Rõ ràng chúng ta không thể xem xét **độ tương tự (Similar)** giữa các đối tượng bằng cảm tính được

Các phương pháp tính toán độ tương tự

Chúng ta cần một cơ sở toán học cụ thể để xác định **độ tương tự** và đại lượng này được gọi là **khoảng cách**:

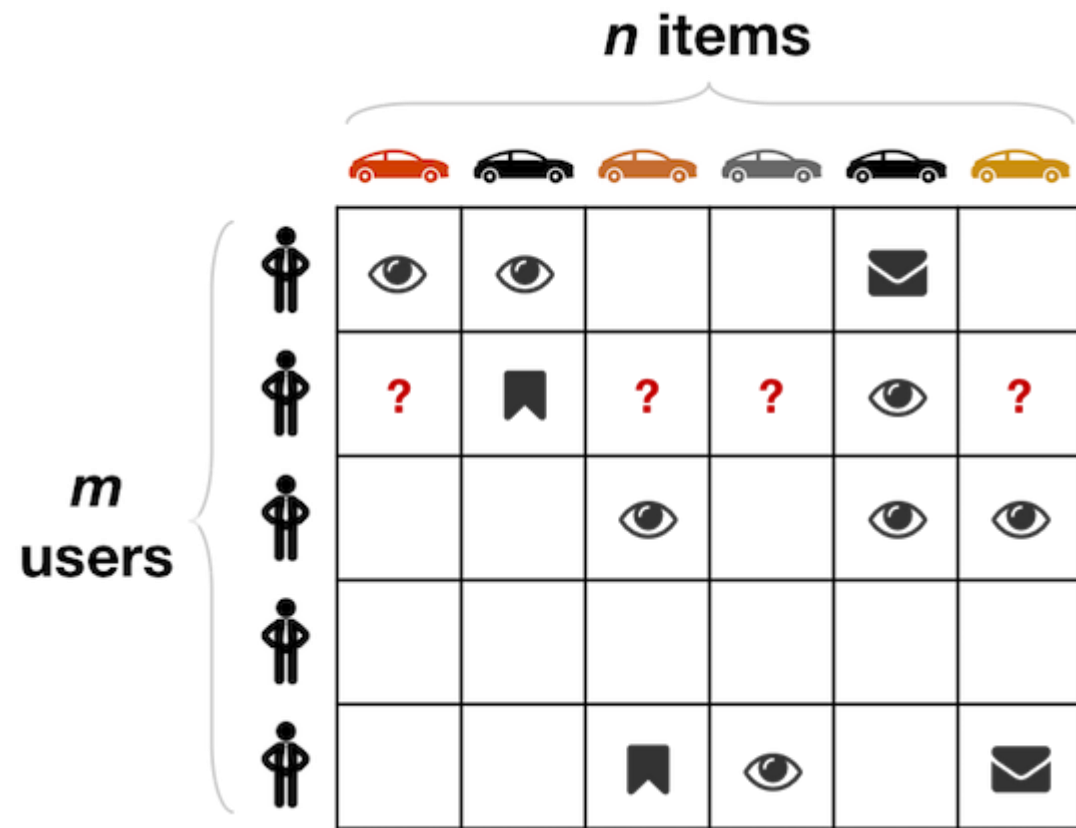
- Khoảng cách **càng nhỏ** => càng gần nhau => độ tương tự **càng lớn**
- Khoảng cách **càng lớn** => Càng xa nhau => độ tương tự **càng nhỏ**

Chúng ta có thể hiểu **độ đo tương tự** giống như **ngịch đảo của khoảng cách**. Sử dụng khoảng cách để tính toán đại lượng này.

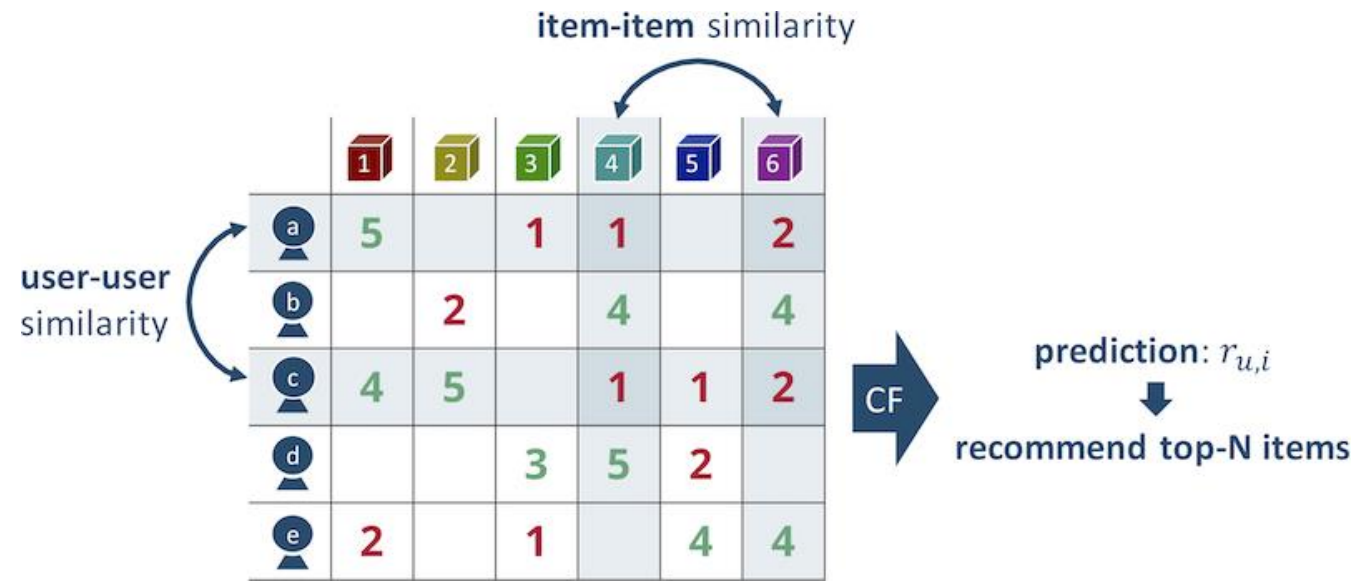


Các phương pháp tính toán độ tương tự

Khoảng cách được tính toán dựa trên ma trận *users - items*



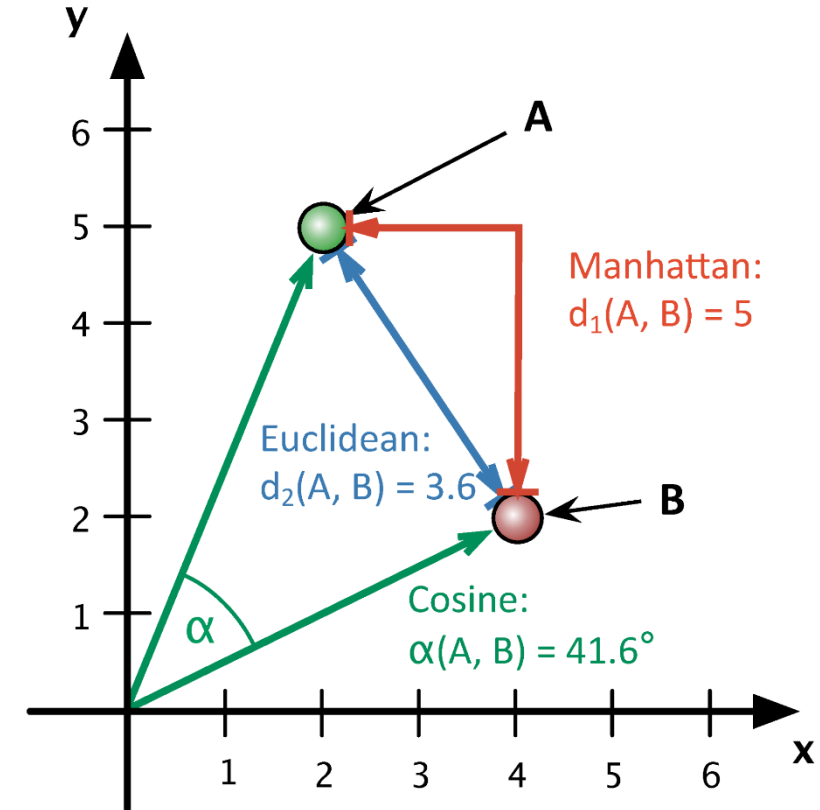
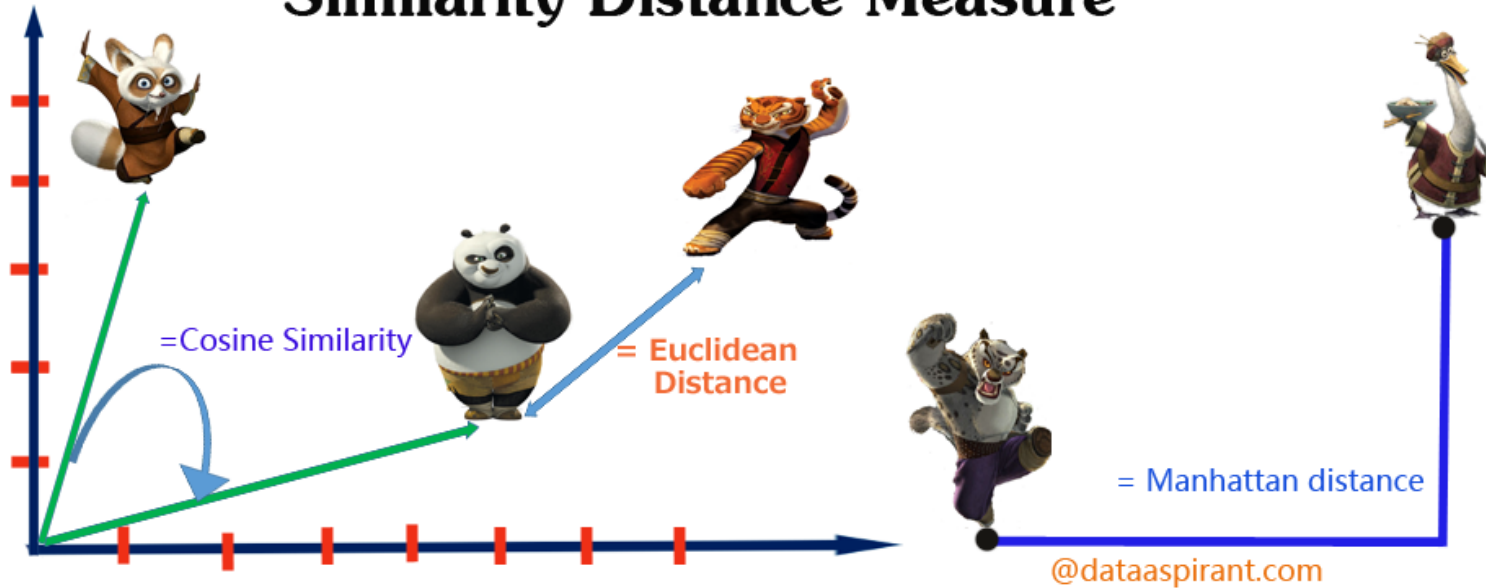
A		✓	✗	✓	✓
B			✓	✗	✗
C		✓	✓	✗	
D		✗		✓	
E		✓	✓	?	✗



Các phương pháp tính toán độ tương tự

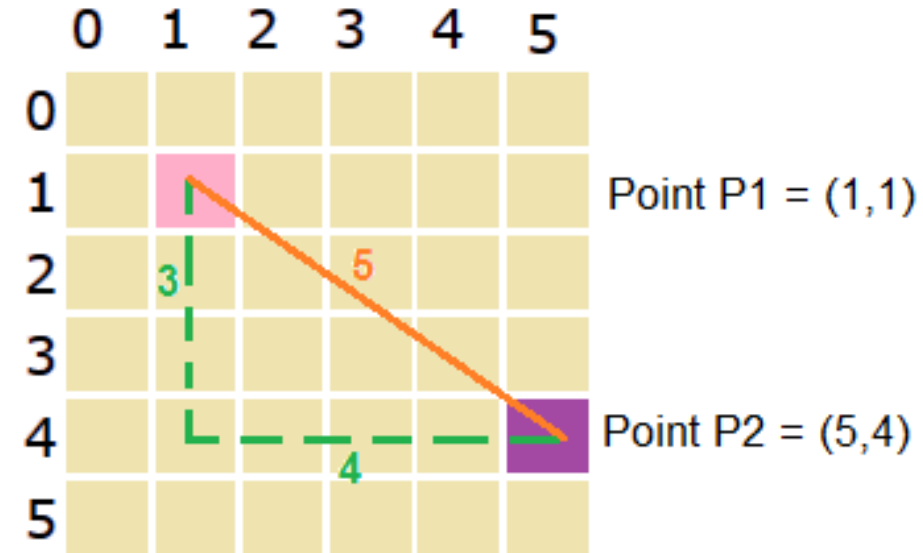
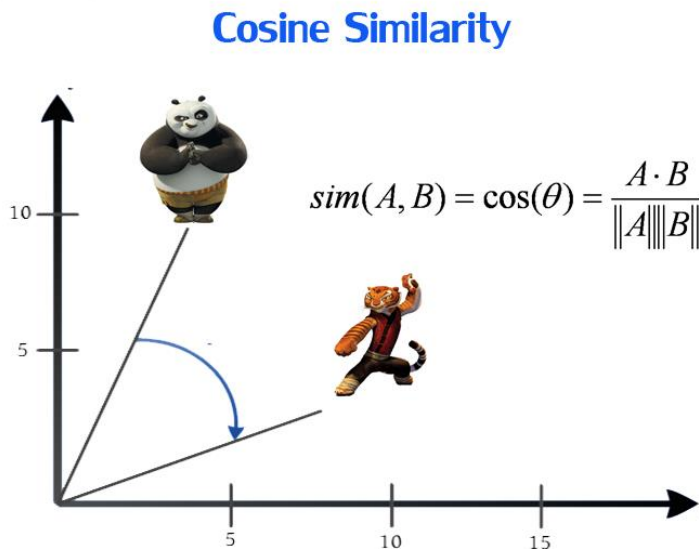
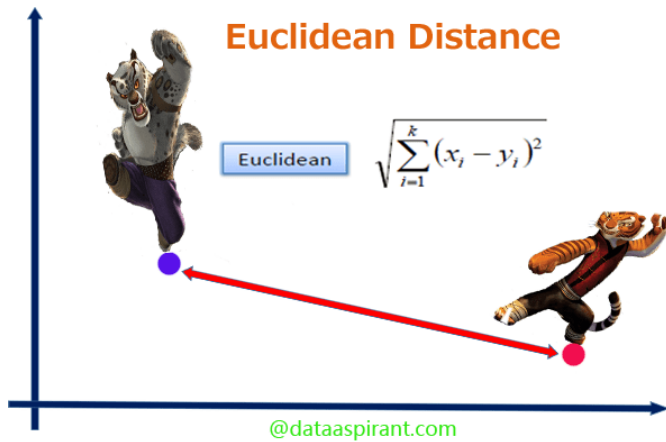
Một số phương pháp tính khoảng cách:

Similarity Distance Measure



Các phương pháp tính toán độ tương tự

Một số phương pháp tính khoảng cách:



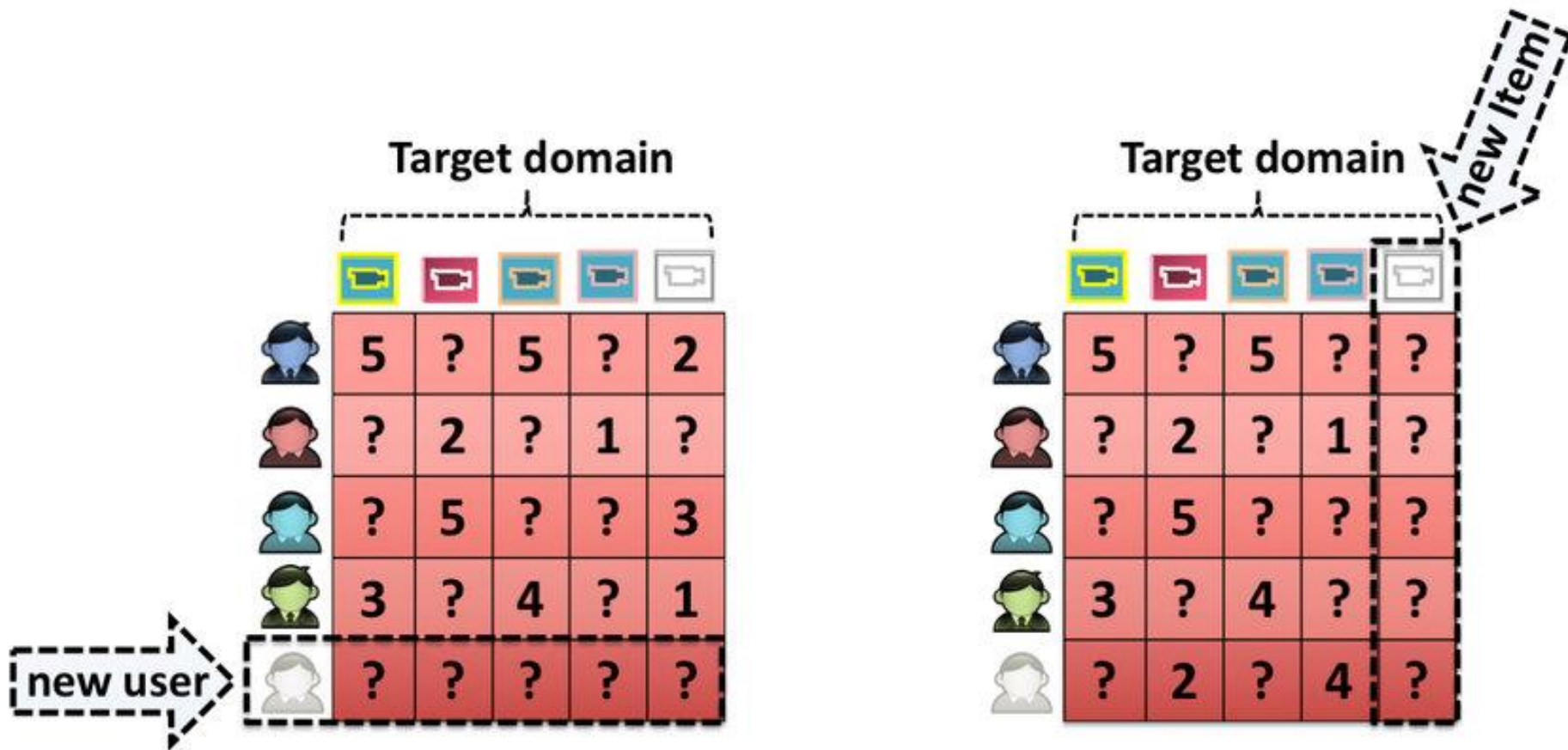
Euclidean distance = $\sqrt{(5-1)^2 + (4-1)^2} = 5$

Manhattan distance = $|5-1| + |4-1| = 7$

4. Thách thức khi xây dựng hệ thống gợi ý

Thách thức

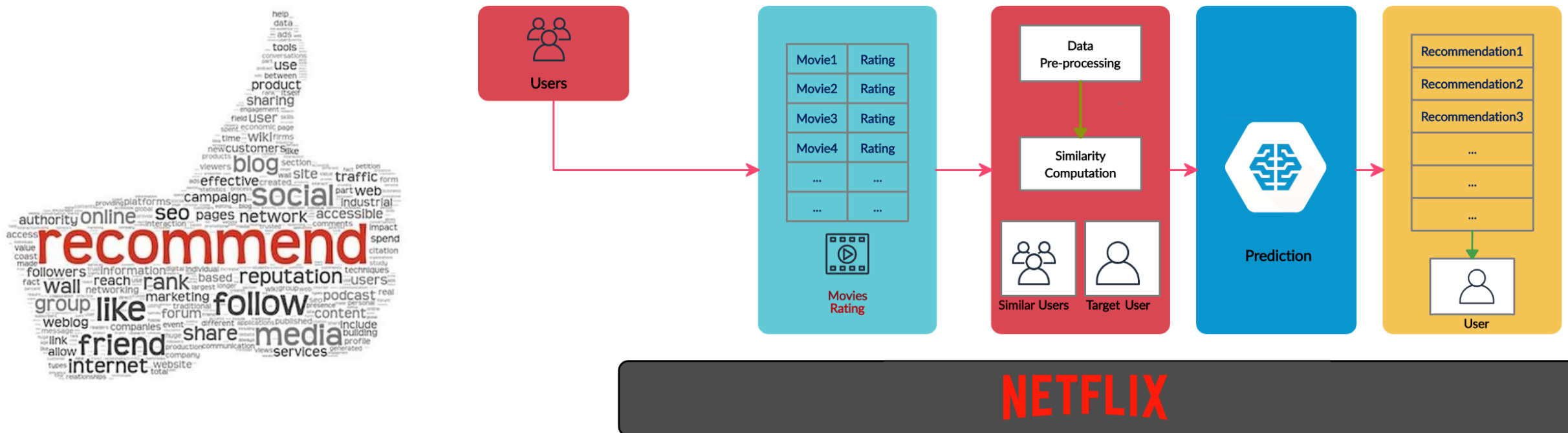
Một trong những thách thức với hệ thống gợi ý đó là vấn đề người dùng mới (new user) – item mới (new item), còn gọi là “cold-start problem”



5. Ví dụ demo hệ thống gợi ý

Giới thiệu bài toán

Xây dựng hệ thống gợi ý phim!





Thank you!