memory-based collaborative filtering in the item view

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | I1 | I2 | I3 | I4 |
| U1 | \* | 1 | 2 | \* |
| U2 | 2 | \* | 4 | \* |
| U3 | 2 | 2 | \* | \* |
| U4 | 5 | 1 | 2 | \* |
| U5 | 2 | 2 | 4 | \* |
| U6 | \* | \* | \* | \* |

Idea= Tu van cho 1 nguoi dung u 1 san pham i neu u da thich nhung san pham tuong doi voi i trong qua khu.

Offline:

-Tinh **do tuong dong** cua tat ca item voi nhau.

SIM(I1,I2)=[(2-2)\*(2-2) + (2-2.5)\*(3-2.5)]:[sqrt((2-2)^2 +(2-2.5)^2) x sqrt((2-2)^2 +(3-2.5)^2)]….

-Sắp xếp:

I1🡨I2;I3;I4

I2🡨I3;I1;I4

I3🡨I2;I4;I1

I4🡸I1,I2,I3

Online:

Active user u1

Ru1i1?.

-Tìm k=2 item láng giềng của i1 mà u1 có đánh giá= I2; I3

Ru1i1 = [**Sim(i1,i2)\*(**Ru1i2-TBu1) + **Sim(i1,i3)\*(**Ru1i3-TBu1)]:[ Sim(i1,i2) + Sim(i1,i3)] +TBU1

So sánh **cộng tác dựa trên láng giềng hướng user và cộng tác dựa trên láng giềng hướng item**

-Chi phí tính toán: số lượng item sẽ ít hơn số lượng user 🡺 hướng item-based tiết kiệm hơn

-Độ đa đạng của danh sách tư vấn: hướng user-based đa đang hướng item-based 🡪 ít nhàm chán.

-Tình diễn giải kết quả tư vấn:

-Item-based: do bạn thích sản phẩm A, sản phẩm B nên tôi dự đoán được bạn thích sản phẩm C 🡺 dễ dàng diễn giải.

-User-based: Những người láng giềng với bạn thích nên tôi tư vấn cho bạn sản phẩm A 🡺 khó diễn giải.

-Tính ổn định của khối offline: user-based << item-based.

Đánh giá điểm mạnh và điểm yếu của **công tac dua tren lang gieng**

Điểm mạnh: đơn giản và gần với con người.

Điểm yếu: -khó triển khai trong hệ thống lớn (đầu tư nhiều về mặt phần cứng: scalability)

* Tính thưa của ma trận user-item (sparsity 🡪 khó tính độ tương đồng
* Tính thưa của ma trận user-item 🡪 khó tổng hợp đánh giá
* Sản phẩm mới --> không chạy được quá trình tư vấn cho sản phẩm đó.
* Người dùng mới 🡪 không chạy được quá trình tư vấn cho người dùng đó.