

1 Giới thiệu

2 Mục tiêu của hệ thống gợi ý

Trước khi bàn luận về mục tiêu của hệ thống gợi ý, chúng ta sẽ nói về một số mô hình khác nhau của hệ thống gợi ý có thể được xây dựng. Có hai mô hình chính:

- Prediction version (phiên bản dự đoán): Cách tiếp cận đầu tiên dùng để dự đoán giá trị rating cho một cặp user-item. Giả sử chúng ta có một tập dữ liệu về mức độ quan tâm của m người dùng (user) lên n sản phẩm (item), tương ứng với một ma trận $m \times n$, trong đó phần tử ở hàng i , cột j thể hiện giá trị rating của user thứ i lên item thứ j . Việc cần làm của chúng ta là lấp đầy các giá trị còn trống trong ma trận, hay nói cách khác, chúng ta sẽ dự đoán giá trị rating của user lên items mà user đó chưa rate. Bài toán này còn được gọi là bài toán hoàn thành ma trận, bởi vì ma trận ban đầu mà chúng ta có được là ma trận chưa đầy đủ giá trị, những giá trị còn lại sẽ được dự đoán bằng cách sử dụng các thuật toán.
- Ranking version (phiên bản xếp hạng): Trong thực tế, thực sự chúng ta không cần thiết phải dự đoán tất cả các giá trị ratings của user lên items để đưa ra gợi ý cho user. Thay vào đó, chúng ta chỉ cần gợi ý top-k items phù hợp nhất cho user, hoặc xác định top-k users phù hợp nhất với item. Thông thường, việc gợi ý k items cho user là phổ biến hơn, mặc dù phương pháp để giải quyết hai vấn đề này là hoàn toàn tương tự nhau.

Trong mô hình thứ hai, giá trị ratings dự đoán được là không quan trọng. Mô hình thứ nhất phổ biến hơn trong thực tế, bởi vì mô hình thứ hai có thể được giải quyết thông qua việc xử lý mô hình thứ nhất. Tuy nhiên, trong nhiều trường hợp, sẽ là dễ dàng hơn và tự nhiên hơn nếu thiết kế và xây dựng mô hình thứ hai một cách trực tiếp.

Trong kinh doanh, việc tăng doanh số bán hàng hay tăng lợi nhuận là nhiệm vụ quan trọng nhất của hệ thống gợi ý. Bằng cách sử dụng các thuật toán, hệ thống gợi ý mang những items thích hợp đến sự chú ý của users, điều này làm tăng doanh thu và lợi nhuận cho người kinh doanh. Để thực hiện được nhiệm vụ này, hệ thống gợi ý phải có một số mục tiêu hoạt động và tiêu chí kỹ thuật như sau:

- Relevance (Tính phù hợp): Một mục tiêu hoạt động hiển nhiên nhất của hệ thống gợi ý là gợi ý items thích hợp với users, do users thường mua những sản phẩm mà họ cảm thấy thú vị và phù hợp. Mặc dù sự thích hợp là mục tiêu chính của hệ thống gợi ý, nhưng nếu chỉ có mỗi mục tiêu này thì vẫn là chưa đủ, vì vậy chúng ta sẽ nói đến một số mục tiêu "hạng 2" sau đây, mặc dù nó không quan trọng bằng Tính phù hợp, nhưng nó vẫn mang một ý nghĩa quan trọng và ảnh hưởng khá nhiều đến một hệ thống gợi ý được xem là tốt.
- Novelty (Tính mới lạ): Hệ thống gợi ý thực sự có hiệu quả khi nó gợi ý những items mà users chưa từng thấy trong quá khứ. Ví dụ, một bộ phim phổ biến với một thể loại quen thuộc sẽ hiếm khi gây ấn tượng cho user. Việc lặp đi lặp lại gợi ý những items phổ biến cũng có thể dẫn tới sự giảm sút trong sự đa dạng trong việc "bán" items.

3 Phân loại một số hệ thống gợi ý phổ biến

3.1 Collaborative Filtering Models

3.2 Content-Based Recommender Systems

3.3 Knowledge-Based Recommender Systems

3.4 Demographic Recommender Systems

3.5 Hybrid and Ensemble-Based Recommender Systems

3.6 Evaluation of Recommender Systems