**Nội dung tóm tắt tuần 1**

**Đề Tài: Xây dựng website đọc sách theo sở thích**

Đặt vấn đề:

Trong thời đại công nghệ đang dần phổ biến trên mọi ngõ ngách, cùng với sự bùng nổ các phương tiện giải trí đã khiến cho giới trẻ dần quên đi thói quen đọc sách. Vì vậy website được xây dựng để thúc đẩy việc đọc sách và cung cấp các công cụ công nghệ để đáp ứng các nhu cầu liên quan tới việc đọc sách.

Hướng triển khai:

* Tập trung vào khảo sát các yêu cầu, mô tả của 1 trang web đọc sách. (thu thập các dữ liệu cần thiết)
* Tìm hiểu khái niệm chung về hệ thống gợi ý, các công nghệ để lập trình web phù hợp
* Đi sâu vào các nhóm thuật toán về hệ thống gợi ý, và thiết kế cơ sở dữ liệu
* Thiết kế và phát triển web.
* Thử nghiệm và đánh giá hệ thống.

- Tìm hiểu về Collaborative Filtering (CF), Content-Based Filtering (CBF), và Hybrid Recommendation Systems: là ba phương pháp chính được sử dụng trong các hệ thống gợi ý để cung cấp các sản phẩm hoặc dịch vụ phù hợp với người dùng dựa trên sở thích và hành vi của họ.

* **Collaborative Filtering (CF):** Phương pháp này dựa trên ý tưởng rằng những người dùng có sở thích tương tự sẽ thích những sản phẩm tương tự. [CF sử dụng dữ liệu đánh giá từ nhiều người dùng để dự đoán sở thích của một người dùng cụ thể](https://www.hindawi.com/journals/sp/2021/7427409/).
* **Content-Based Filtering (CBF):** CBF khuyến nghị các sản phẩm dựa trên đặc điểm của chúng và sự tương đồng với các sản phẩm mà người dùng đã thích trước đây. [Nó tập trung vào nội dung của các mặt hàng và sử dụng thông tin mô tả sản phẩm để đưa ra gợi ý](https://www.hindawi.com/journals/sp/2021/7427409/).
* **Hybrid Recommendation System**: Hệ thống gợi ý lai kết hợp cả hai phương pháp CF và CBF để tận dụng ưu điểm của cả hai, nhằm cải thiện chất lượng gợi ý và giảm bớt nhược điểm của từng phương pháp khi đứng một mình. [Hệ thống lai có thể sử dụng nhiều kỹ thuật khác nhau như đánh trọng số cho các gợi ý từ CF và CBF, hoặc sử dụng một mô hình dự đoán chung cho cả hai loại dữ liệu](https://www.hindawi.com/journals/sp/2021/7427409/).

- So sánh giữa **Collaborative Filtering (CF)**, **Content-Based Filtering (CBF)**, và **Hybrid Recommendation Systems**:

1. **CF**:
   * **Ưu điểm**:
     + Khả năng phát hiện sở thích mới không dựa trên nội dung.
     + Cải thiện theo thời gian khi có thêm dữ liệu từ người dùng.
   * **Nhược điểm**:
     + Cần lượng dữ liệu lớn để hiệu quả (vấn đề khởi đầu lạnh).
     + Khó khăn trong việc xử lý các sản phẩm mới không có đánh giá.
2. **CBF**:
   * **Ưu điểm**:
     + Khuyến nghị dựa trên nội dung cụ thể của sản phẩm, tốt cho các mặt hàng mới.
     + Dễ dàng giải thích lý do khuyến nghị.
   * **Nhược điểm**:
     + Giới hạn trong việc khám phá sở thích mới của người dùng.
     + Cần phân tích nội dung chi tiết, có thể phức tạp.
3. **Hybrid**:
   * **Ưu điểm**:
     + Kết hợp ưu điểm của cả CF và CBF, giảm thiểu nhược điểm của từng phương pháp.
     + Linh hoạt trong việc xử lý dữ liệu và cải thiện chất lượng gợi ý.
   * **Nhược điểm**:
     + Có thể phức tạp hơn để triển khai và tối ưu hóa.
     + Cần cân nhắc kỹ lưỡng khi kết hợp các phương pháp để tránh xung đột.

Tìm hiểu thuật toán kNN : Thuật toán kNN (k-Nearest Neighbors): Đây là một phương pháp lọc cộng tác, nơi hệ thống sẽ tìm kiếm k người dùng gần nhất (tức là có sở thích tương tự) và dùng dữ liệu của họ để đưa ra đề xuất.

- Lợi ích :

* Đơn giản và trực quan
* Không cần mô hình phức tạp
* [Cập](https://aurigait.com/blog/recommendation-system-using-knn/) nhập dễ dàng
* [Phù](https://aurigait.com/blog/recommendation-system-using-knn/) hợp với các loại dữ liệu khác nhau

So sánh giữa KNN và Recommender System:

|  |  |  |
| --- | --- | --- |
|  | KNN | Recommender System |
| Mục tiêu web | Khuyến khích khám phá:  + Nếu mục tiêu chính là khuyến khích người dùng khám phá những cuốn sách mới, KNN có thể là lựa chọn tốt hơn.  + KNN có thể đề xuất những cuốn sách tương tự với những cuốn sách mà người dùng đã đọc hoặc đánh giá cao, giúp họ khám phá những tác giả hoặc thể loại mới. | Cá nhân hóa trải nghiệm:  + Nếu mục tiêu chính là cá nhân hóa trải nghiệm đọc sách cho người dùng, Recommender System có thể phù hợp hơn.  + Recommender System có thể phân tích sở thích và hành vi đọc sách của người dùng để đưa ra những đề xuất phù hợp với nhu cầu và gu đọc của họ. |
| Dữ liệu | **Dữ liệu ít:**  **+** hoạt động hiệu quả với lượng dữ liệu người dùng ít hơn so với Recommender System  + chỉ cần dữ liệu về những cuốn sách mà người dùng đã đọc hoặc đánh giá cao để đưa ra đề xuất. | Dữ liệu nhiều:  + hoạt động hiệu quả hơn khi có lượng dữ liệu người dùng lớn.  + phân tích nhiều dữ liệu hơn, chẳng hạn như lịch sử đọc sách, đánh giá, xếp hạng, thời gian đọc, v.v., để đưa ra những đề xuất chính xác hơn. |
| Chi Phí | Rẻ hơn để triển khai và vận hành vì không yêu cầu nhiều tài nguyên | Đắt đỏ hơn để triển khai và vận hành, đặc biệt là khi sử dụng các thuật toán phức tạp và dữ liệu lớn. |
|  |  |  |
|  |  |  |

=> Theo như yêu cầu của web thì nhóm em thấy thuật toán Recommender System phù hợp hơn. Vì web sẽ đề xuất những cuốn sách theo sở thích của user và cá nhân hóa để tăng tính trải nghiệm của người đọc.