Case study report (HW part 1)
Goodreads.com

—

Deepak Ravishankar (UNI: dr2998)

Link(https://www.goodreads.com/)

Introduction

Goodreads.com is a website which recommends you books based on books you have read and let's you read other people's reviews and write your own.

After the signup process, the user is presented the home page where the user can be a part of communities and book recommendations are given on this page. Due to the company being acquired by amazon and the integration with kindle, to update the progress of your books is as simple as taping a button. The books are already arranged into 3 categories titled "Books read", "Currently reading" and "want to read. The user can also create custom shelves to arrange his books, and you get shelf specific recommendations.

Personalization and recommendation elements

The site uses a mixture of personalization and recommendation. Currently there are simple elements which are personalized like, current reading list and progress of books you are reading (synced using kindle or manually), amount of reading challenge completed and books you want to read. The user can also create custom shelves and based on the shelves the user is recommended more books to add to that shelf. There is a mix of both, where the user is given recommendations based on the books he has read or wants to read. This helps improve discovery for the user from their catalog of more than 395 million books.

The website offers a five star rating system with the ability for the user to write a written review as well. Also due to collaboration with amazon the website does have access to the content of all the books as well. The recommendation and ratings of amazon books and good reads are still kept separate though. (based on a statement from amazon)

The website uses recommendation for other parts of the service as well such as such as user recommendation of groups to join based on various things like "Groups reading your To-Read books" etc.

Personalization and recommendation Assessment

The website does a very good job at recommending new books and keeping the user motivated and coming back through personalized challenges. The personalization aspect works very well due to the kindle sync feature which is seamless. The recommendations are very relevant (from personal experience I have read books I discovered purely through goodreads). The recommendations also are not completely in the same domain and the website tries to recommend new authors which you might like.

Development concerns

There are a multitude of concerns that were discussed that apply to the website like :

User facing concerns :
Interpretability: the website does this by showing the user why a particular book is recommended like "based on your interest in philosophy" or "based on your reading list etc."

Rich get richer: This is a huge problem as usually a few popular books are highly read and usually have multiple reviews and other books have none. The service does recommend books with very few reviews to no reviews as well. This might have been done by recommending books based on their content as well and not just by comparing reviews. As book content is

readily available to the service due to their collaboration with amazon this might be a viable strategy

Serendipity and diversity: The website urges the user to go through the award winning books in other genres to help the user widen his taste and discover more books and might help you explore more.

Engineering concerns:
Cold start: The service during the sign-up process, asks you for your topics of interest and asks you to rate books that you have already read. This is a good workaround to the cold start problem and helps you get some initial data on the user. Even during the rating process, based on the books you have rated till now , the algorithms tries to find other books you might have read, making the discovery of books already read much smoother.

Batch vs real-time processing: The service mostly does batch processing for the recommendation carousels as it shows when the results were last calculated. All the recommendations for the user are also not calculated at the same time as all the carousels have different timestamps as to when they were calculated.

Scalability: This must be a huge problem for the site as it has more than 20 million members and more than 50 million reviews. Batch processing definitely helps here as you do not need to calculate frequently for the same user. Also as people usually read at a slow pace the data updation is not as common as other use cases as even a prolific user might read one book per week which can be easily processed in a batch.

Adding new features: This must definitely be a important part of their algorithm. As postulated above they must be using content based recommendations and as more and more breakthroughs in NLP happen everyday, there would be a need to use these new features and update their model.

Changes and additions

I am currently thoroughly enjoying using this service, but the only thing i don't like is that either your entire library is visible to all your friends or you can hide your entire library. You might not want your friends to see some of your rating and books read (in categories like erotica or topics like coping with sexual abuse etc.) and this might lead to a problem of a user not sharing all his interests or have a false rating on book just because other people are notified of their ratings. So I would add a feature to hide certain book rating from being personally identifiable.

Important Aspects to replicate

The very simple sync feature with a kindle might not be the most innovative is definitely the most useful feature of personalization and any competing product will have to have a similar feature. This makes updating your book progress a smooth experience. Also the gamification of reading of setting a goal for the user is also a personalization feature which makes user visit the website again to update his goal number of book per year. The recommendation engine's ability to recommend less reviewed books would also be a very important feature to replicate in a competing product

Conclusion

The website does a good job at covering all the required aspects of making a good recommendation engine. The personalization aspects of website make it a custom experience for each user and keep the user engaged and must be a part of any new and similar product.
■ ■ ■