Text Information Systems – Technology Review

Collaborative Filtering Applications and Challenges

Introduction

Everyday people use the internet to satisfy their information needs, their short-term needs are satisfied through browsers while their long-term information needs require extra curation. This extra curation is provided through recommender systems [1]. Recommender systems keep track of user information to they can provide suggestions to users about what they may enjoy. Recommender systems have become a ubiquitous part of the internet and greatly influence how people experience content on the web. In this context it becomes obvious that recommender systems can have strong and unexpected effects on people.

Applications

A recommender system is essentially a filtering system [1]. There are two ways that filtering can be done. The first way is content based filtering where items are suggested based on their similarity to other items the user liked in the past [1] the second way, which will be discussed in this review, is collaborative filtering where users are clustered groups and suggestions are made based on what other members of this cluster or group like [1]. An advantage of this collaborative filtering is that the system does not need to have knowledge of the content of items it is suggesting. However, many recommender systems use both types of filtering to make recommendations. Collaborative filtering is both useful and influential, it has changed how people receive content on the internet and ultimately controls how people interact with the digital landscape.

Some of the most influential online businesses use collaborative filtering. These are companies like Amazon, Facebook, Twitter, LinkedIn, Spotify, Google News, Last.fm [2] and more. In this context it becomes a little more apparent that collaborative filtering influences what you watch, listen to, and even buy and read. In this context collaborative filtering is integral to a great deal of online activities.

Challenges

Collaborative filtering is widely used and as such it is not surprising that it can have can have profound effects on the users whose experiences are shaped by it. Unfortunately, the effects are not always positive things like finding your new favorite song or being able to immediately see the news that interests you.

A question that comes up when looking at ethics in collaborative filtering is what kinds of items should and should not be suggested to users. Indeed, even if a user will like an item it does not mean the recommender system should suggest it to them [4]. Additionally, some users may find some types of content upsetting even if said content is socially considered. Furthermore, while users may generally have similar interests and therefore have similar taste In other items this does not necessarily mean that ethically their preferences will align, this means it could be advisable to have an ethical filter after the collaborative filtering stage to ensure that users are not receiving upsetting or inappropriate content. One solution to this challenge could be to create a database of ethical content which could be determined using the laws and social and cultural norms of a country, this database would then in a way act a filter for recommendations [3] this would serve to simply remove all socially deemed inappropriate content from items that can be recommended. Another solution is to take a more customizable approach where a user can adjust "ethical" filters, so items are separated out based on the users specific ethical preferences [4]. This second approach could solve the issue of some users finding things upsetting that are commonly not considered so. Perhaps, these two approaches could be used together to first remove content that is not appropriate for any user and then do the filtering on a more personal level to adjust for personal preference.

Collaborative filtering groups similar users to make recommendations, this means that users will likely only be exposed to items like what they already consume since users similar to them, or rather in their group, would consume similar items; this can have the unexpected effect of re-enforcing extreme user views. In other words collaborative filtering can make an echo chamber where users are only exposed to views that mirror their own, a so called "filter bubble" [5] with self-reinforcing biases which can prove to be detrimental to balanced public debates [5]. This is a cycle which is hard to break since collaborative filtering creates the above-mentioned conditions by virtue of how it works, however a method to mitigate these issue would be to not exclusively use collaborative filtering to give suggestions. Combing collaborative filtering with another method would give the best of both worlds, letting user see things they are comfortable with and likely to enjoy but also other items that can expand their tastes [6], this mixture of different methods would help to find a balance in the explore exploit trade off.

Recommender systems can push users in a specific direction by restricting the variety of content that the user is exposed to [7], resulting in the user forming a preference they otherwise may not. While this filtering can be helpful to allow the users to spend less times making decisions by reducing ther number of items for users to sift through, the result can also be morally questionable if the intent was to persuade or push the user into having a certain preference [7]. A way to overcome this challenge could be to allow users to see why they are recommended certain items. Giving an explanation would help by allowing users to decide what aligns with their personal preferences and sense of self. However, giving this explanation could be problematic as they may identify categories for users that the users themselves do not agree with which could be damaging to their personal identity. When categorization does not follow commonly accepted social categories the users' experience of personal identity can be disrupted [8]. Another challenge is that the system does not have a complete view of the user but is instead constantly changing the representation of the user based on their feedback and interaction with the system (implicit feedback).

Recommender systems are based on observed data and because of this it is easy for them to be affected by bias that is present in the training data, due to bias present in the real world. When trying to remove bias from algorithms it is common to remove sensitive features such as age, gender, and race. While removing these features can help reduce the unfairness it is often not enough [9]. It is worth noting that collaborative filtering does not directly consider features but instead infers features from user behavior [9]. However, when performing recommendation, it is important to remember that sensitive features such as age, race and gender likely do play a role in user preferences [9] and therefore it would be illogical not to use them. The two primary sources of bias in recommender systems come from problematic patterns of data collection. Specifically observational bias is an issue which occurs when recommendation to specific groups of user group creates feedback loops [6]. Another source is general population imbalance where the data being used for the system reflects the existing social model and therefore expressing bias towards some groups [6]. In this context to overcome these challenges the data should be collected fairly, with possible inequalities being examined before the time of collection and after collection. The data should then be normalized to adjust for any bias. Additionally, there are several measures that can be utilized to compare fairness [6] such as *value fairness* which measures inconsistency in signed estimation error across the user types*, absolute fairness* which measures inconsistency in absolute estimation error and *non-parity unfairness* measure based on the regularization term and *underestimation unfairness* which measures inconsistency in how much the predictions underestimate the true ratings [9]. Using these measures can help the system creators to understand when unfairness is present.

Conclusion

In conclusion collaborative filtering has many applications and is very useful both to companies and individual users when creating a recommender system there are many issues that should be considered. Indeed, a key challenged that is often present where machine learning is involved is how we can control bias that is unintentionally present in the system from data the system was training on and critically understanding the implication the new technology will have on society. Collaborative filtering greatly controls the experience users have online which can ultimately affect their online behavior and even behaviors that reach beyond the screen. That being said when the challenges are considered, and solutions are found to overcome these challenges the technology of collaborative filtering is groundbreaking and can improve the daily online lives of millions of people helping them get what they need faster and more efficiently.

Citations

1. Zhai, C., 2021. Recommender Systems- part 2.
2. Techcrunch.com. 2021. TechCrunch is part of the Yahoo family of brands. [online] Available at: <https://techcrunch.com/2015/09/28/the-evolving-landscape-of-recommendation-systems/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAALooxHi2uDIi5hK4WiXoxitYItBcZdNHJfKlXuxiZERMu8KwbJjAqQC84v5-Oregw5jzNNR8pOJMKlaAfgTO7VcAJpqFO3zbThLY5qZaYt68WsbLgUXqXJxUPFU7Kew3iSdTzAbvMtzq1Jq7p8QQrXYo89IuIP0sEla_ZAbLK4Cy#:~:text=Collaborative%20Filtering&text=Comp

anies%20that%20employ%20this%20model,Google%20News%20and%20Last.fm.>
[Accessed 3 November 2021].

3. Souali, K., El Afia, A. and Faizi, R., 2011, April. An automatic ethical-based recommender system for e-commerce. In 2011 International Conference on Multimedia Computing and Systems (pp. 1-4). IEEE.
4. Tang, T.Y. and Winoto, P., 2016. I should not recommend it to you even if you will like it: the ethics of recommender systems. New Review of Hypermedia and Multimedia, 22(1-2), pp.111-138.
5. Reviglio, U., 2017, November. Serendipity by design? How to turn from diversity exposure to diversity experience to face filter bubbles in social media. In International Conference on Internet Science (pp. 281-300). Springer, Cham.
6. Milano, S., Taddeo, M. and Floridi, L., 2020. Recommender systems and their ethical challenges. AI & SOCIETY, 35(4), pp.957-967.
7. Burr, C., Cristianini, N. and Ladyman, J., 2018. An analysis of the interaction between intelligent software agents and human users. Minds and machines, 28(4), pp.735-774.
8. De Vries, K., 2010. Identity, profiling algorithms and a world of ambient intelligence. *Ethics and information technology*, *12*(1), pp.71-85.
9. Yao, S. and Huang, B., 2017. Beyond parity: Fairness objectives for collaborative filtering. *arXiv preprint arXiv:1705.08804*.