

Collaborative Filtering and Innovation: A Case Study from Steam Recommendations

Coen D. Needell

Tue, Mar 16, 2021

Introduction

The modern world is flooded with options; which TV show to watch, which video games to play, which espresso machine to buy. E-commerce is characterized by its quantity. This doesn't, however, imply that e-commerce platforms are inundated with low quality goods. Truly fantastic media is released every year. 2019 saw the release of a game *Disco Elysium* which has been lauded as one of the best CRPGs (Computer Role Playing Games) of all time, even though the so-called golden age of the genre is long past (Metacritic, n.d.). 2019 also saw the release of 8033 games on Steam, PC gaming's most used distribution platform (Statista, n.d.). Television has experienced a similar trend, with some calling this the Golden Age of Television Writing. However, these trends present a problem. Clearly not all 8033 games that were released onto Steam in 2019 were brilliant, genre-redefining masterworks like *Disco Elysium*, *Outer Wilds*, and *Untitled Goose Game*. Players don't have the money and energy to invest in games that they aren't going to enjoy. They need someone to separate the wheat from the chaff for them.

Netflix famously tried to address this problem with a million-dollar competition to design a new content discovery algorithm. While the current algorithm is proprietary, from prospectus documents, the Netflix prize algorithms, and press releases, it can be surmised that Netflix's recommendations come from a mixture of clustering and collaborative filtering, incorporating both user data and content data. Netflix lives and dies by its recommendation algorithm, so it's become fairly good over the years. Steam, however, is in a very different situation. While Steam puts a lot of resources into making better content discovery tools, it's less focused. Steam has, by my count, eight different content discovery systems. The "discovery queue," the "interactive recommender," the "play next shelf," the "deep dive" (deprecated), the "community recommends" page, the "sale explorer," the "navigator," and the front page has an infinite scroll-style recommendation system. All of these employ some aspect of personalization or machine learning inference.

Steam is, by its very nature, overwhelming to the consumer. This is a side-effect of the fact that there's a very low barrier to entry to get on Steam. Your product just needs to not break any laws, be your right to sell, not threaten the security of computers it's installed on, and not be pornography to get in¹ (Steamworks, n.d.). It's part of Steam's core philosophy that anyone using their platform succeeds or fails under their own power, rather than because Steam themselves decided that their game should succeed². For contrast, Steam's biggest competitor, the Epic Games Store, for example, has highly restrictive barriers to entry, but also guarantees sales (Epic buys the first X copies and resells them, taking the loss themselves if they don't sell) and promotes new releases heavily. Steam leans into the quantity approach, but it also has plenty of tools to help a player sift through the thousands of games. In addition to the AI driven techniques above, there are also human driven systems, like a curator system, where anyone can create a curated collection of Steam games and people who find they have similar tastes can get recommendations that way. Every game on steam has an open review space, where anyone who has bought and played the game can leave a review with a binary recommendation tag. These systems are known to have their problems as well, such as instances where a game is inundated with mobs of bad reviews in retaliation to something unrelated the developer did. These concerns are outside the scope of this discussion, however, we are going to look at the problems with automated recommendation.

Statistical methods for recommendation are a great way to get information on large groups of people. They're even quite good at grouping people together. They also make fantastically accurate predictions in well-trodden ground,

¹This bar is even lower than you think, as we will discover.

²Valve does sell some of their own products on Steam, but they don't really push them very hard. I can't remember the last time I saw any Valve games come up in my Steam recommendations.

but when you go off the beaten path, their accuracy falters. Statistical prediction systems of any kind are playing a game of averages and expectations. More sophisticated methods perform better on new situations because they're better at recognizing patterns, but they still can be tripped up when presented with patterns they don't know how to parse. Ultimately, they use the patterns they do recognize, and make a prediction based on the average of things they've seen exhibit similar patterns. So here's the problem, game developers, and the media industry in general, are in the business of doing things that have never been done. It is, in a sense, their jobs to go off the beaten path. Most recommendation algorithms, and indeed the technique that this discussion looks at, are based on the assumption that if two users liked the same things in the past, then they'll like the same things in the future. This is not a bad assumption in general, but there are cases where it fails. For a thought experiment, consider two groups of gamers, the Strategy Buffs and the Casuals. Casuals generally prefer clever gameplay, engaging, but not overwhelmingly deep. They tend to gravitate towards simulation and puzzle games. The Strategy Buffs, however, love deep tactics, they like the feeling of outsmarting an opponent. There is some overlap in the games these groups enjoy, but not a lot. Now let's say a game like *Wargroove* is released. It's a tactics game, but it doesn't run too deep, it's about outsmarting your opponent, but it doesn't have long, grueling campaigns, it can be played in thirty-minute chunks while watching TV or between getting home from work and cooking dinner. It appeals to some Strategy Buffs because it has some innovations in the Medieval-Tactics subgenre, and it appeals to Casuals because it's easy to pick up, and it has cute graphics. Now imagine a user who plays *Wargroove*, and likes it, and wants to use the recommendation system to find more games like it. What is the optimal result? This is one of many quirks of the recommendation problem, and one in particular that, if ignored, can punish game developers for innovating.

The Recommendation Problem

From the point of view of the consumer, there are N options of varying quality. Each of them is unique. In addition to their quality, there's also a sense of how good of a match that option is for the consumer. Some people like action, some people like to play online, some people prefer strategic thinking. In addition to people's general preferences, they also have in-the-moment preferences.

References

- Metacritic. n.d. "Disco Elysium." *Metacritic*. <https://www.metacritic.com/game/pc/disco-elysium>.
- Statista. n.d. "Number of Games Released on Steam 2020." *Statista*. <https://www.statista.com/statistics/552623/number-games-released-steam/>.
- Steamworks. n.d. "Steamworks Partner Program." <https://partner.steamgames.com/steamdirect>.