Niche Cinema: The Impact of Netflix's Recommendation System

Ori Spector

Stanford University

March 9th, 2022

Slides: Niche Cinema

Introduction

You are what you watch. Open up your Netflix account and look at your home page. Look at the selections available to you. Ask yourself: Does this content describe me? You might have been oblivious to the fact that everything you see is part of your watching personality. These options are highly tailored to each individual based on the data that has been fed to Netflix's algorithmic system time and time again. At the core of this "phenomenon" stands recommendation systems. Recommendation systems use millions of data points to construct accurate prediction results to target the preferences of individuals. However, these big data products are nothing without the individuals who use them.

This paper will analyze the Netflix recommendation system and dive into the intricacies of how the algorithm uses data to present highly tailored content that best fits one's preferences. In other words, this research aims to explore the question: What is the effect of Netflix's recommendation system as it uses people as products? Is it homogenizing our cinematic taste? Is it clumping individuals into content-based bubbles and hurting the diversification of content watched? To answer these questions, we need to define specific terms, understand the underlying foundation of recommendation systems, look at research case studies, and examine the overall impacts and rationales.

People as Products

Before we can jump into the research, it is critical to understand the underlying principles of the consumer-product model commonly utilized in the big-data world. However, to truly understand the consumer-product model, we must first understand the core structure that builds it. The exact intricacies of the technology are outside the scope of this research paper, but the fundamental idea is essential to comprehend. The underlying structure of recommendation systems is machine learning. *Deep learning* is defined as "a machine learning technique that teaches computers to do what comes naturally to humans: learn by example (Mathworks, 1)". The most common implementation of deep learning techniques used by big-data companies is a method known as neural networks. This method is composed of a series of algorithms that endeavor to recognize underlying relationships in a set of data through a process that mimics how the human brain operates (MathWorks, 1). In this sense, neural networks refer to either organic or artificial neurons. To keep it in layman's terms, Netflix uses large quantities of data to compose deep learning algorithms that constantly learn and output responses that are highly tailored to the inputs they receive. At a high level, the recommendation system output says: "this is what you want to watch." Of course, it is more complex than this in reality, but the point is that data drives the output of the system. The more data, the more accurate the response

Neil Hunt et al. provide research that emphasizes the complex layers in Netflix's recommendation output framework in his paper titled: *The Netflix Recommender System: Algorithms, Business Value, and Innovation* from *ACM Journals*. Categorically, the research reports six crucial focuses of the system: personalized feed, personalized video rank, trending content, video-video similarity, continue-watching page, and page generation (Hunt, 4-6). These factors contribute to how the user reacts to the product visually and implicitly. Netflix determines what is best in all these categories for the subscribers by analyzing what leads to the most engagement.

How does this information have anything to do with a consumer-product model? Netflix boasts an astounding 200M users; they constantly collect subscriber data and use the deep learning algorithms above to build their recommendation system. Every time a user selects something to watch, more data is retrieved by Netflix. This information feeds right into their machine learning system, creating a positive feedback loop that gives users even more accurate suggestions (Chong, 1). These outputs are highly tailored to the individual's preferences, maximizing one's engagement and creating a cyclical pattern of users feeding them data, improving the system continuously. Explicitly, the definition of the consumer-user model for this paper is as follows: the concept of using "user" data to power a product—using people as products (Tabora, 1). Thus, we can see that Netflix falls victim to using this method to make their product more engaging and superior. This leads us to the quintessential question of the research. How does Netflix's recommendation system, which utilizes the consumer-user model, impact us?

Netflix and the Effects of the Recommendation System

What is the goal of Netflix's recommendation system? As Netflix's research team puts it, the goal is "providing our members with personalized suggestions to reduce the amount of time and frustration to find great content to watch." When one opens Netflix, the page of movies and television shows one sees is a visual feed representing what Netflix recommends that an individual watch. In an honors thesis by Stanford graduate Julianna Yonis, titled "*You Are What You Watch*," she describes the rationale for and against the recommendation system Netflix employs. She presents a common argument from proponents of the algorithmic recommendation system of Netflix by stating that "they make diversified content economically feasible. (Yonis, 71)" Essentially, Netflix provides thousands of options to stream content compared with other avenues of watching audiovisual content (such as movie theaters or Cable networks) that are much more costly to consumers. Additionally, supporters of personalization highlight the advantageous nature of deep learning algorithms, readily making tailored recommendations for them to watch. However, most Netflix consumers do not think about how the recommendation system influences or impacts them "under-the-hood." How does Netflix's algorithmic system affect us? Two key observations have climbed into the spotlight throughout my research to answer this question: the homogenization effect and anti-diversification clumping.

The Homogenization Effect

One of the Netflix recommendation system's significant impacts is *the homogenization effect.* The homogenization effect is the concept that Netflix's recommendation system can target specific content to a broad audience of people, which ends up narrowing our individual preferences. Consequently, this makes us, the users, watch the same content. In essence, our cinematic tastes are homogenized. This repercussion is rooted in technology exploiting data to learn about our tendencies and maximize engagement. This realization comes from a series of experiments on the Netflix recommendation system by scholar Niko Pajovic from *Sage Journal*.

The case study by Pajovic sheds light on a revelation that Netflix's recommendation system has led to the homogenization of our cinematic taste. On day one of his experiment, he starts with three different Netflix profiles to represent three different types of people who might be users: a die-hard sports fan, a hopeless romantic, and a culture snob (Pajovic, 7). The die-hard sports fanatic streams sports-related content such as the Michael Jordan documentary and *Rocky*. The hopeless romantic watches films like *The Notebook* and *Crazy, Stupid, Love*. Finally, the culture snob views critically acclaimed cinematic "masterpieces," including the *Godfather* and *Citizen Kane*. After a week of viewing appropriate content for each user, the genre rows such as 'Exciting Movies' and 'Familiar Favorites' had been highly tailored to each profile (Pajovic, 8).

Curiously, personalization was even visible within the 'Popular on Netflix' row, which sounds objective and impersonal, but turned out to align with user interests (Pajovic, 9).

Why is the case study critical? Netflix produced a new original show called *Outer Banks,* which follows high school students uncovering hidden treasure. The show itself was not directly aligned with the personas' tastes. However, Netflix's recommendation algorithm marketed it to each user in an appealing way tailored to their taste. For the sports fan, the cover picture was the two male leads with surfboards. The cover art was a close-up of the main romantic couple embracing the hopeless romantic persona. Finally, the cover art was an abstract still of the main character behind a map for the cinema snob (Pajovic, 12). We have three people with distinct tastes being marketed in ways tailored to their desires, increasing the chance they watch the same series or movies. This study is a prime example of an implicit impact the algorithmic system of recommendations has on unknowing users.

Why is this effect occurring?

We can see the homogenization effect in action, but why is this trend emerging? We need to dive deeper into Netflix's motives to answer why this impact is apparent. Netflix, on a superficial level, desires to promote the best content personalized to each user, but at its core as a business, Netflix is influenced by monetary gains. Thus, they intend to preserve an extensive foundation of subscribers. To do so, Netflix has two primary directions it can follow to address its goals. The first is to distribute and produce content exceptionally tailored to each consumer's distinct preferences. The second is to alter subscribers' tastes to align with the existing content.

The algorithmic system is constructed to match the latter method. This is a consequence led predominantly by the cost of producing and licensing content. Looking at statistics provided by Statista, we can see a graph of Netflix's content expenditure over the years, and we can derive trends in the streaming service industry. 46% of Netflix's content costs come from their in-house original content production, increasing year over year. These costs amount to over $11 billion of a total $25 billion budget as of 2020. This is an enormously high cost of production. Thus, it is unsustainable and unrealistic to cater to every user's taste at such high fees. Thus, the homogenization effect is a result of Netflix's monetary incentives.

Overall Impacts

It is now clear why the homogenization effect is present, but how does this affect users? The homogenization effect emphasizes that Netflix invests in projects that their algorithmic recommendation system is trained to promote. This content will engage the most users, retaining their majority mainstream user foundation. By focusing on the mainstream, a significant result is the loss of niche cinema. Niche Cinema is the category of films and television that caters to the interests of the non-majority and non-mainstream audiences. Statistically, less content is available on Netflix as their total available options of streaming have drastically reduced by thousands (nearly 20%) over the last five years (Statista). This is important because users need to realize this impact in a broader sense. The audience and consumers of content are being shown fewer options. Notably, we are receiving fewer recommendations of older and niche categories in exchange for new and mainstream choices. This is an immediate result of broadening the mainstream content favoring Netflix's economic goals. Additionally, all users are impacted as the existing content is marketed to a broader range of people than fit their exact preferences. Thus, over time this means the recommendation system will push niche viewers' cinematic taste to fit into the algorithms bubble, essentially eradicating broader cinematic views. Moreover, users make biased decisions when choosing Netflix content to watch. This is an implicit effect as the recommended content is biased in favor of Netflix's agenda.

Silver linings in the Homogenization Effect

Importantly, as a reader and potentially a subscriber of Netflix's product, one at the very least knows about an implicit bias that exists in the decision-making process of selecting content. The loss of niche cinema caused by the homogenization effect is now apparent. However, we can also analyze the potential upsides of the homogenization effect. Statistically, Netflix does output a high volume of critically acclaimed movies. Critically acclaimed productions are films or shows that received awards and praise from industry professionals or critics (YourDictionary). Explicitly, Netflix produced 109 Oscars nominated projects and 518 Emmy-nominated shows since 2018, per Statista research. These numbers are the highest by a landslide amount compared to other streaming services and production companies. Thus, we might be streaming the same content, but potentially the content we watch is more significant or praised by professionals in cinema. Additionally, the mainstream contains most people; therefore, potentially only a smaller percentage of users have their niche preferences affected. Lastly, people are naturally conversational; watching the same content makes us more socially aligned and can relate to or discuss more cinematic topics together.

Anti-Diversification Clumping

The second widespread impact of the recommendation system is *Anti-Diversification Clumping.* This concept states that we share similar tastes as other users, forming preference-adjacent groups. Then once we join a cluster, we are recommended to only content that fits our profile. This result essentially shuts us out from seeing a wider area of available content, hurting the diversification of our watch history. This occurs predominantly due to the algorithmic system trapping us in a bubble. To emphasize the point, journalist Sameer Chhabra from MobileSyrup, a Canadian news source for technical articles, interviews Todd Yellin, Netflix's vice president of product innovation. Yellin talks about an example of clumping in action. He begins by stating that "approximately 80 percent of subscribers trust and follow the recommendations of the algorithm" (Chhabra, 1). In other words, people click on Netflix's recommended content feed 80% of the time. In practice, this means that what one sees is what one gets. Additionally, Yellin explains how Netflix creates their data clumps, which cluster users into inevitable watch bubbles. He states,

> "We refer to genres as wrappers because they are a good descriptor of a mobile or TV show, but they aren't necessarily complete…The algorithm also categorizes content based on 'thousands of additional qualifiers,' like mood, aesthetic, and pace. (Chhabra, 1)"

To illustrate this point further, he gives an example from Marvel's *Jessica Jones*, a superhero show in the genre of a detective thriller.

> "Yellin explained that, to Netflix, *Jessica Jones* is a dark crime drama like *Making a Murderer*; it has a strong female lead like *Orange Is the New Black*; and it features a sharp sense of humour like *Master of None*. If you've watched one of those three shows, then you've also probably received a recommendation to watch *Jessica Jones*. Likewise, viewers of *Jessica Jones have* most likely also received recommendations to watch those other three shows. The same is true for shows like Marvel's *Daredevil* and *The Defenders*. (Chhabra, 1)"

From this, we can see a visual representation of a web of possible suggestions that stem from viewing interests. This is a positive feedback loop meaning the user will become trapped in the algorithm's recommended shows. However, the recommendation algorithm finds even more connections from user activity. This is a crucial realization because by grouping people into specific genre bubbles, niche genres or non-generic plots become less targeted and desired by Netflix.

Additionally, anti-diversification clumping can be seen in a case study by WBR Insights. They point out some poignant consequences and controversies that result from preference-based clumping. Specifically, content clumping can be founded on a subscriber's race, ethnicity, and gender. Although Netflix does not explicitly ask for someone's demographic, the algorithm can make assumptions based on user data. The case study concentrated on the cover artwork personalization the subscriber sees for each video. After implementing this new personalization thumbnail algorithm, numerous African-American users reported feeling racially profiled and targeted. Specifically, the images they were seeing "were racially and ethnically driven, and were often misrepresentative of the actual cast of the movie. (WBR Insights, 1)" For instance, one of the African-American individuals in the study reported that the thumbnail he saw for the romantic comedy film *Love Actually* featured "British black actor Chiwetel Ejiofor, who plays a minor role in the film, and Keira Knightley. (WBR Insights, 1)" However, many non-black subscribers saw a thumbnail featuring the main stars of the film: Hugh Grant, Emma Thompson, and Colin Firth. (WBR Insights, 1)" This phenomenon occurred with other films, such as the murder mystery series *The Good Cop*. This presents another example of how the recommendation system engenders user clumping. Thus, as an African-American user selects a film with a targeted black Actor or Actress in the cover image, the algorithmic system will only provide more visually similar titles.

Thus, the diversity of content consumers' view is reduced due to the user clumping phenomenon. Moreover, since people are likely to keep clicking on titles based on the image, the user data reinforces the positive feedback loop. Therefore, there is clear evidence that algorithmic clumping exists on both a genre-rooted and profile-based level.

Why is this effect occurring?

We see the primarily implicit impact the Netflix recommendation system creates by clumping users into viewing bubbles from this anti-diversification effect. At the core of this, Stanford Ph.D. candidate Julianna Yonis points out, "the system limits user autonomy by collecting data without the user's active participation, constraining the user's behavior, and manufacturing the user's desires. (Yonis, iii)". This allows Netflix to biasedly fit less content into more users' "preferences." This is a subjective incentive that aligns once again with Netflix's monetary motivations. Another major factor that drives Netflix's decision to implement the algorithm in this manner is because of the result from their research: "visual elements were the biggest influencing factor when a viewer was deciding which title to watch, comprising 82% of the focus" (Netflix, 1). By grouping people together based on data, Netflix can better "fit" their content into what the user thinks they desire. While the content recommendation is accurately tailored to individual subscribers, there is a gradual shift in the decision-making process. The subliminal phenomenon is an effect of the algorithm hidden but momentous in action. Users are placed into buckets and are made to accept them.

Conclusion and Takeaways

We have studied Netflix's system to recommend items to its subscribers. We have seen two massive observations that result from the technology employed in the homogenization effect and anti-diversification clumping. These effects produce implicit and explicit impacts for the user when they decide what to watch. These are likely results of the economic motives of Netflix. Netflix desires to shift audiences towards a mainstream cinematic taste. This is intrinsically against their core vision, which promises content personalization to their user base. This is a result that personalization and the economics of streaming are on opposite sides. Thus, we see the Netflix recommendation system is learning to evolve to support Netflix and not the subscribers who power the machine (consumer-product model). Thus, the most vital impact of the algorithmic system is that Netflix is constructing a user that behaves how they need them to, as opposed to how they naturally want to. This engenders an illusion of user taste preferences. Sadly, the monetary gains of mainstream consumption of media warrant that Netflix's primary concentration cannot be aligned with niche audiences. This results in the impact that mainstream content will be the target of content production, and less-generic, eccentric pieces of cinema will be slowly eradicated.

.

Works Cited

Basilico, Justin. *Recommender Systems in Industry: A Netflix Case ... - Springer*. https://link.springer.com/chapter/10.1007%2F978-1-4899-7637-6_11.

Chong, David. "Deep Dive into Netflix's Recommender System." *Medium*, Towards Data Science, 24 Sept. 2021, https://towardsdatascience.com/deep-dive-into-netflixs-recommender-system-341806ae3b48.

Chhabra, Sameer. "Netflix Says 80 Percent of Watched Content Is Based on Algorithmic Recommendations." *MobileSyrup*, 21 Aug. 2017, https://mobilesyrup.com/2017/08/22/80-percent-netflix-shows-discovered-recommendation/.

"Critically-Acclaimed Meaning." *Critically-Acclaimed Meaning | Best 2 Definitions of Critically-Acclaimed*, https://www.yourdictionary.com/critically-acclaimed.

Hunt, Niel, et al. "The Netflix Recommender System: Algorithms, Business Value, and Innovation: ACM Transactions on Management Information Systems: Vol 6, No 4." *ACM Transactions on Management Information Systems*, 1 Jan. 2016, https://dl.acm.org/doi/10.1145/2843948.

Navarro, José Gabriel. "Oscar Nominations per Film Studio 2021." *Statista*, 12 Aug. 2021, https://www.statista.com/statistics/681263/oscars-nominations-per-film-studio/.

"Netflix Research." *Netflix Research*, https://research.netflix.com/research-area/recommendations.

Pajkovic, Niko. "Algorithms and Taste-Making ... - Journals.sagepub.com." *Sage Journals*, https://journals.sagepub.com/doi/full/10.1177/13548565211014464.

Stoll, Julia. "Netflix: Content Assets 2020." *Statista*, 18 Mar. 2021, https://www.statista.com/statistics/881673/netflix-content-assets/.

Stoll, Julia. "Netflix Oscar Nominations 2022." *Statista*, 9 Feb. 2022, https://www.statista.com/statistics/1231555/netflix-oscar-nominations/.

Stoll, Julia. "Emmy Awards: Netflix Nominations and Wins 2013-2021." *Statista*, 25 Oct. 2021, https://www.statista.com/statistics/324190/nominations-winners-netflix-original-programs-emmy-awards-usa/.

Tabora, Vincent. “In Big Data, the Consumer Is the Product.” *Medium*, The Startup, 1 Nov. 2019, https://medium.com/swlh/in-big-data-the-consumer-is-the-product-ad9bf0423d9a.

WBRInsights. “The Netflix Effect: How Netflix's Commitment to Extreme Personalization Is Impacting Customer Expectations.” *NGCX 2022*, https://nextgencx.wbresearch.com/blog/netflix-effect-commitment-extreme-personalization-customer-expectations.

“What Is Deep Learning?: How It Works, Techniques & Applications.” *How It Works, Techniques & Applications - MATLAB & Simulink*, https://www.mathworks.com/discovery/deep-learning.html.

“Why Am I Seeing This?” *New America*, https://www.newamerica.org/oti/reports/why-am-i-seeing-this/case-study-netflix/.

Yonis, Julianna (2021). You Are What You Watch: The Netflix Recommendation System and the Rise of the Algorithmic Culture Industry. Stanford Digital Repository. Available at: https://purl.stanford.edu/zz193vr5979