Is the use of YouTube's AI recommendation system actually beneficial for the user?

Marcus Malak
4451570

Jevon Ravenberg
4896270

Tim Verburg
4783891

Xinyu Wang
5454522

I. Introduction

YouTube has become an increasingly popular platform over the past years. At the moment it is the most popular video-sharing platform with over 2 billion active users [1]. To keep users engaged YouTube makes use of a personalized AI recommendation system to provide relevant new videos to active users. These recommendations are based on a variety of metrics like geo-location, watch history and gender. Even though YouTube has a search function easily accessible to users, 70 percent of the user watch time was driven by the AI recommendation algorithm as stated by YouTube Chief Product Officer Neal Mohan at CES 2018 [2]. With that large a user base, the use of this algorithm has a large societal impact. Besides it being convenient for the users we need to think about the consequences of bias and oversights in a system like that being used. This leads to our research question:

Is the use of YouTube's AI recommendation system actually beneficial for the user?

The research question will be evaluated using two ethical frameworks, the consequentialist and the duty framework. We will provide the general structure of the YouTube AI recommendation system and the inherent consequences of its design. In addition we will discuss the availability of misinformative content and the resulting effects on the user base and the existence of so called "filter bubbles". These areas were considered to be the most influential factors and were therefore chosen to be evaluated. More can be said about the YouTube recommendation system and its effects, to manage the scope this report will be limited to discussing the previously mentioned topics.

II. Analysis and evaluation of AI technology

A. How does YouTube's recommender AI work (Jevon)

The general working principle of the YouTube recommendation system is described by Covington et al. [3], figure 1 shows a schematic of its overall architecture. It is important to note that this source is quite old and there is no real way to truly know all the details about the current implementation of YouTube's AI recommendation system, however this section will serve as at least a basic overview of the implementation. The system is based on Tensorflow [4], an end-to-end machine
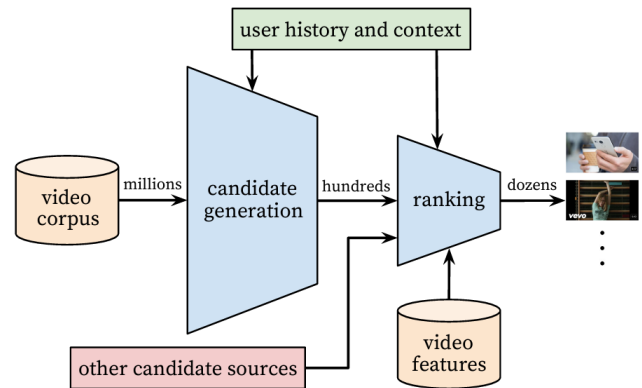

Fig. 1. Recommendation system architecture [3]

learning tool that was developed by Google and made open-source. The system is comprised of two deep neural network stages, the first is for candidate generation, the second for ranking.

The candidate generation network takes features based on the user's YouTube activity history and context as an input to filter out a subset of hundreds of candidate recommendations from the millions of videos on YouTube. The features derived from user data include watch history, search history, geographic region, device, gender, age and logged-in state. The video features are crafted from video meta-data, i.e. views, age, and video content signals, i.e. tags [5].

The ranking network takes these candidates and potentially candidates from other sources and assigns them a score based on features describing the user, context and also the videos. Because the amount of videos being processed here is orders of magnitude lower than in the previous stage, much more features describing the video and the user's relationship to the video are used. The candidates are then ranked on their score and the highest ranking videos are presented to the user in order.

In the training process of the algorithm, which can either be done offline or online, Covington et al. [3] describes steps taken to remove certain biases. Training samples are generated from all YouTube watches, not just watches from produced recommendations. Moreover the amount of training samples generated per user is fixed, which prevents the group of highly active users from dominating training performance

(loss). Furthermore the age of a training sample, so how long a certain video has been in the training window, is taken into account in order to remove neural networks' inherent bias towards old samples. This means more recent videos within the training window have higher likelihood of being recommended.

The algorithm is constantly trying to improve its performance and for that it needs some metric to judge whether a recommendation was accurate or not. Covington et al. [3] mentions specifically that the metric used is not explicit user feedback, i.e. thumbs up/down, but watch time as explicit feedback is far more sparse. A more recent research paper [5] however states that this explicit feedback is also used, but we think it is very likely that the watch time is still the predominant factor. This means the algorithm recommends videos based on the predicted rating value and not per se on whether the video fits the user.

YouTube makes money via advertisements embedded into videos, subscriptions and premium content and aids Google in data collection which could generate even more revenue [6]. This first point would imply that YouTube has a vested interest in maximizing viewer watch time as this would maximize revenue from advertisements.

B. Ethical framework

1) The Consequentialist Framework:: The consequentialist framework focuses on the outcome of a possible action, taking into account those directly and indirectly affected. Evaluating purely the outcome might not always be good as it relies on a prediction of the outcome, which might turn out worse than anticipated.

2) The Duty Framework:: The duty frameworks focuses on the implied duties and obligations that a person or group of people has in a certain situation. The ethically correct decision is defined by doing the right thing based on a persons duties. This framework expects the same from every person in a situation, therefore promoting a certain equality.

TODO: Further explain framework

C. The consequence of maximizing engagement

Due to the nature of YouTube's business model where the goal is to keep you engaged as long as possible, people naturally spend a long time watching video after video. The AI keeps recommending interesting videos and without knowing you find yourself spending lots of time on YouTube. In some cases this can actually lead to addiction as a person is craving new impulses and is never satisfied. An article from the Guardian [7], shows an example of a person who got in the habit of watching so many videos that it severely impacted his daily life. We need to argue if it's responsible from YouTube to have such a system in place knowing that it could lead to addicted people, thus having a negative impact on their lives. Addicted users mean more money for YouTube so it is actually beneficial for them that people fall into this habit. Looking at it from the duty framework we have to argue that YouTube has a responsibility towards their user

base to care for their well being. There is no evidence of them actively limiting the watch time of users so they are actually pushing people towards addiction on purpose. It has been shown that addiction to internet in general leads to depression and social isolation [8]. From the duty perspective we think YouTube should take responsibility towards their users. According to the YouTube user statistics of 2022 [1] there are a staggering 2 billion active users. On this scale the social impact is huge and it can't be neglected that YouTube needs to be held accountable for the social impact of their platform.

Another perspective we can use is the consequentialist framework applied on the actual use of the AI. The algorithm provides personalized recommendations which normally looks convenient. This can be true in some cases where one watches one or two videos on a topic you are actually looking for. As soon as we go past that we get to the point where a user is unknowingly being held on the website by feeding them interesting impulses. This can eventually lead to addiction which is a direct consequence of the YouTube AI. The intended use of the technology seems convenient as a user always gets recommended engaging content based on their interests. Unfortunately due to the nature of YouTube's business model the technology can have a negative impact on the user.

D. Feeding people controversial information. Shape user preferences and guide choices.

Misinformation is a huge problem in modern society as evident by the increase in vaccine hesitancy [9]. A paper by Hussein and Juneja et al. [10] explores the claims that YouTube promotes misinformation. What they found is that, within their chosen areas, new accounts are unlikely to be presented with promoting misinformation videos [10], with the exception of the chemtrials category. Neither age nor geo-location seemed to have a significant impact on misinformation unlike the watch history [10], which is further supported by Ledwich and Zaitsev [11], who also concluded that most of the recommendations presented to a user within an extreme group were based on the watch history. Hussein and Juneja et al. noted that it is evident that since YouTube's efforts to combat certain misinformative topics like medical misinformation, the percentage of recommendations and search results are significantly less than that of lesser known misinformation topic like chemtrails [10].

Viewing the user of YouTube's AI recommendation algorithm from a consequentialist perspective reinforces the importance of reliable information sources. Hussein and Juneja et al. found that having the account gender set to male increases the amount of misinformative videos recommended to them, where in 4 out of the 8 categories evaluated, the male account was recommended significantly higher amount misinformative videos even when watching neutral videos [10]. As further mentioned in the paper, they stressed the societal implications of this as they noted that a study revealed that the percentage of male anti-vaxxers was higher then for females,

thus the recommendation algorithm might be reinforcing the conspiracy beliefs. A paper by Spinelli and Crovella [12] noted the algorithm's tendency to promote more extreme and less trustworthy content based on whether the user was logged in, not logged in, in private mode, or even using a Tor browser [13] (a browser designed around anonymity). Contrary to this, the paper by Ledwich and Zaitsev [11] explored the political landscape of the YouTube recommendations in which they found that the algorithm favours mainstream content over the extremes. It even goes as far as to recommend users in these extremes videos of the lesser extremes [11] which could be interpreted as a form of deradicalisation.

As viewed from the duty framework, YouTube should limit the amount of misinformation and actively recommend misinformed users debunking videos. YouTube's selective efforts to combat the misinformation are indeed effective [10], however it is worrying as the topics picked by YouTube are those presented in popular media. The lesser known misinformative topics might be just as dangerous while being just as popular. YouTube should therefore further expand its effort to combat misinformation based on the colossal effect it has on society.

E. filter bubble: loss of diversity

The term filter bubble was coined by internet activist Eli Pariser circa 2010 [14]. the 'filter bubble' is a persistent concept which suggests we no longer encounter a balanced and healthy information diet, but only see information that targets our established interests and reinforces our existing worldviews [15].
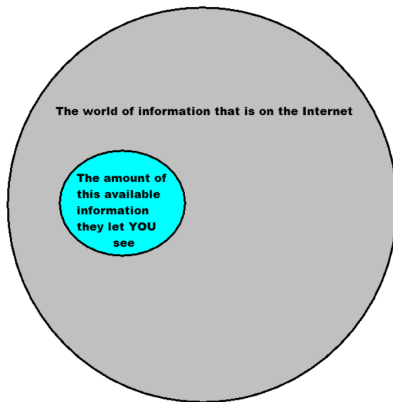

Fig. 2. filter bubble hide info

As discussed in previous section, YouTube's AI algorithm is built not to assist consumers but maximise user attention [16]. To achieve that, users are surrounded by similar sources that match our interest and search history. from a consequential perspective, Being in a filter bubble means we will continuously be fed views that are pleasant, familiar, and confirms our beliefs—and because these filters are invisible, we won't know what is being hidden from us. Our past interests will determine what we are exposed to in the future, leaving less room for the unexpected encounters that spark creativity, innovation, and the democratic exchange of ideas.

For instance, YouTube may hide videos from your friends with different viewpoints and you all get stuck in your own bubbles. you do not know what is being recommended to your friends. That makes it difficult for you to have a productive conversation if you don't even know what kind of information your friends are being exposed to. Thus a lack of understanding and an unwillingness to consider opposing viewpoints emerge.

On the other hand, however, the application of AI recommendation system on YouTube seems plausible. As previously mentioned, in 2021, almost 720,000 hours of video were uploaded worldwide to YouTube every single day [17]. The users have to query the search engine on YouTube, which has become the second-most popular search engine after Google [17], to find what they are looking for. However, It costs to maintain a powerful search engine to deal with an extremely large amount of query and users may also suffer from irrelevant search results. According to Friedman's shareholder theory [18], the only social responsibility that organizations have is to satisfy their owners. From this aspect, it is YouTube's duty to attract users by AI recommendation system, which targets the user's preference precisely and feeds interesting content straight to the user.

Although AI recommendation systems provides users with a convenient and accurate access to the "video world", it should not function with bias, which then results in a possible filter bubble. There are growing expectations for corporations to be socially responsible and reliable [19]. Companies should not just undertake strategies that generate higher profits regardless of how much they benefit society. There is already a lot of problematic content on YouTube. Papadamou [20] have showed that young children are not only able, but are likely to encounter disturbing videos. If teenagers get trapped in a filter bubble full of disturbing videos, their thinking and behavior pattern might be negatively and irreversibly affected. YouTube should absolutely take social responsibility to mitigate filter bubble problem and become a real open social media instead of isolating people.

III. Conclusion

(e.g. 300 words)
Re-introduce and answer your research question.
• Clear recommendations based on the analysis and evaluation.
• Include a list of future potential questions that follow from your analysis.
• The section in general logically follows from the preceding sections.

References

[1] "YouTube User Statistics 2022 | Global Media Insight," Mar 2022. [Online; accessed 8. Mar. 2022].

[2] J. E. Solsman, "YouTube's AI is the puppet master over most of what you watch," *CNET*, Jan 2018.

[3] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for youtube recommendations," in *Proceedings of the 10th ACM Conference on Recommender Systems*, (New York, NY, USA), 2016.

[4] "Tensorflow." (Accessed on 03/10/2022).

[5] Z. Zhao, L. Hong, L. Wei, J. Chen, A. Nath, S. Andrews, A. Kumthekar, M. Sathiamoorthy, X. Yi, and E. Chi, "Recommending what video to watch next: A multitask ranking system," in *Proceedings of the 13th ACM Conference on Recommender Systems*, RecSys '19, (New York, NY, USA), p. 43–51, Association for Computing Machinery, 2019.

[6] "How YouTube Makes Money Off Videos," Mar 2022. [Online; accessed 10. Mar. 2022].

[7] G. s. Reporter, "YouTube addiction: binge watching videos became my 'drug of choice'," *The Guardian*, Sep 2021.

[8] M. Shaw and D. W. Black, "Internet Addiction," *CNS Drugs*, vol. 22, pp. 353–365, May 2008.

[9] "Ten threats to global health in 2019," Mar 2022. [Online; accessed 6. Mar. 2022].

[10] E. Hussein, P. Juneja, and T. Mitra, "Measuring misinformation in video search platforms: An audit study on youtube," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–27, 2020.

[11] M. Ledwich and A. Zaitsev, "Algorithmic extremism: Examining youtube's rabbit hole of radicalization," *arXiv preprint arXiv:1912.11211*, 2019.

[12] L. Spinelli and M. Crovella, "How youtube leads privacy-seeking users away from reliable information," in *Adjunct publication of the 28th ACM conference on user modeling, adaptation and personalization*, pp. 244–251, 2020.

[13] "The Tor Project | Privacy & Freedom Online," Feb 2022. [Online; accessed 8. Mar. 2022].

[14] Contributors to Wikimedia projects, "Filter bubble - Wikipedia," Feb 2022. [Online; accessed 9. Mar. 2022].

[15] A. Bruns, "Filter bubble," *Internet Policy Review*, vol. 8, Nov 2019.

[16] "Trustworthiness of YouTube Recommendations: Artificial Intelligence and Bias," Jul 2020. [Online; accessed 9. Mar. 2022].

[17] M. Mohsin, "10 Youtube Statistics That You Need to Know in 2021," *Oberlo*, Jan 2022.

[18] M. Friedman, "The Social Responsibility of Business Is to Increase Its Profits," in *Corporate Ethics and Corporate Governance*, pp. 173–178, Berlin, Germany: Springer, 2007.

[19] K. Fukukawa, *Corporate social responsibility in Asia*. Abingdon, England, UK: Abingdon, UK : Routledge, 2010., 2010.

[20] K. Papadamou, A. Papasavva, S. Zannettou, J. Blackburn, N. Kourtellis, I. Leontiadis, G. Stringhini, and M. Sirivianos, "Disturbed YouTube for Kids: Characterizing and Detecting Inappropriate Videos Targeting Young Children," *ICWSM*, vol. 14, pp. 522–533, May 2020.