

Further Examination of YouTube’s Rabbit-Hole Algorithm



Matthew Moldawsky

1 Background

1.1 Previous Research

For many years, YouTube has been criticized by many mass media outlets for having a “right-leaning” radicalized recommendation algorithm. In other words, they claimed that the algorithm prioritized politically right channels. Last year, a report was made that sought to challenge these claims. Mark Ledwich and Anna Zaitsev [1] compiled over 800 YouTube channels that have over 10,000 subscribers, and more than 30 percent of the content is political. They set out to assess common claims from the media using a data set. They categorized the channels using different tags in order to gauge the impressions from the type of content. They had people watch content from all these channels to assign political labels through unanimous decision from the laborers. From the data that the team collected, they were able to conclude that the algorithm does not recommend content that might contribute to a radicalization of the user base. The data that they collected actually showed that the algorithm leans more toward content that falls within mainstream media. Ledwich even acknowledges the limitations of their method to their research. The research paper finishes with a conclusion that “one cannot proclaim that YouTube’s algorithm . . . is leading users towards more radical content” [1].

M. Moldawsky (✉)
Marist College, Poughkeepsie, NY, USA
e-mail: Matthew.Moldawsky1@marist.edu

1.2 *Other Important Factors and Statements*

Some important factors were not discussed in the previously mentioned paper. For instance, the recommendations for users vary depending on if they are signed in or not which the study did not cover. This caused some criticism of the paper online from various analysts, most notably from Arvind Narayanan [2]. He stated that he did not think it was really feasible to do a quantitative study of the algorithm. Another angle that should be considered is that this “radicalization” is not just about politics. For example, if a user only watches videos about certain video games, then that are mostly all of the content they will be shown. There is a clear relationship between the algorithm, users, and content creators. Radicalization, in this case, is more so the algorithm attempting to prioritize user preferences for content. The problem with this is that certain content paths can lead to extreme versions of the content a user already watches.

In late 2019, the Mozilla Foundation gathered a collection of 28 stories from various users of YouTube in which people fell down a “rabbit-hole” [3]. Lastly, it should be noted that the YouTube algorithm was designed to prioritize certain aspects of video in order to rank it. One of those aspects is engaging content as the algorithm is designed to keep users watching as much as possible. However, according to a dev who worked on the algorithm, “the more outlandish content you make, the more likely it’ll keep people watching, which in turn will make it more likely to be recommended by the algorithm” [4]. This explains why many users could be going down this “rabbit-hole” and could end up being recommended unsavory videos. This then becomes an ethical issue because this type of extreme content could be shown to children depending on the situation. Furthermore, Google, who owns YouTube, has stated previously that they wish to “recommend even more targeted content to users in the interest of increasing engagement” [5]. In other words, Google wants the YouTube platform to become more addicting by increasing the engagement of the platform. Making the platform addicting could be dangerous for children. Figure 1 shows the users across all ages of the YouTube platform.

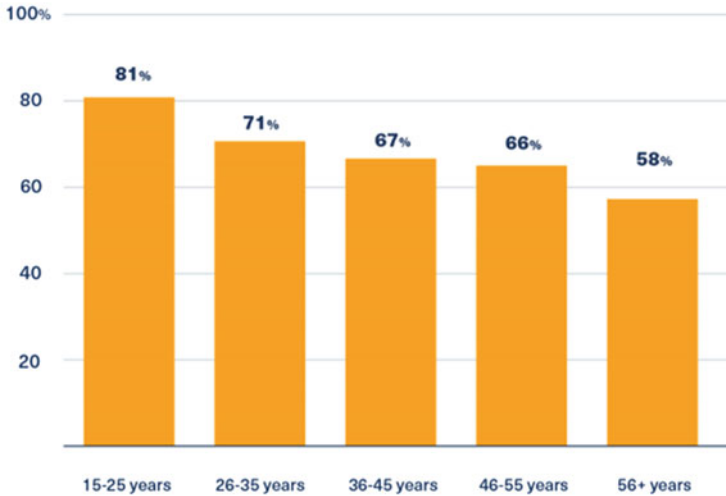
2 **Anecdotes**

2.1 *The Unintended Problems*

The following are a few anecdotes gathered by the Mozilla Foundation and then given to YouTube. Their intention was to show YouTube the problems that could arise from targeted content to increase engagement.

I started searching for “fail videos” where people fall or get a little hurt. I was then presented with a channel that showed dash cam videos from cars. At first it was minor accidents, but later it transitioned into cars blowing up and falling off

Percentage of U.S. internet users who use YouTube as of 3rd quarter 2019, by age group



Source: Statista

Fig. 1 A bar group that shows US users across a wide age range from last year [6]

bridges—videos where people clearly didn’t survive the accident. I felt a little bit sick at that point, and haven’t really sought out that type of content after that.

These terrible videos just keep being recommended to her. She is now restricting her eating and drinking. I heard her downstairs saying “work to eat! work to drink!” I don’t know how I can undo the damage that’s been done to her impressionable mind.

But my recommendations and the sidebar were full of anti-LGBT and similar hateful content. It got to the point where I stopped watching their content and still regretted it, as the recommendations followed me for ages after.

3 Conclusion

Based on the anecdotal evidence and the information given by both YouTube and Google, we can compile a list of the current problems with the YouTube algorithm. The anecdotes show that the algorithm prioritizes a user’s preferences for content. YouTube has told us that the algorithm is designed this way. The anecdotes show why YouTube’s algorithm can be described as a “rabbit-hole.” The more of a

certain type of content you watch and search for, then the more the algorithm will recommend it. Soon after searching for one type of content, you'll be recommended more. If the user clicks on that recommended video, then they will be recommended even more. The problem with this method is that some unintended problems could occur as shown by the anecdotes. The algorithm is meant to get you addicted to YouTube. The claims about radicalization are somewhat true in that users will be recommended more extreme content of what they already watch.

YouTube has already taken steps to adjust the algorithm in order to correct some of the problems many have brought up. The main problem is that the algorithm is too focused on giving a personalized experience to the user. There are some potential solutions that YouTube should investigate in order to make a healthier experience for the user. For example, instead of only recommending videos that the algorithm thinks the user might like, maybe the algorithm could include videos that other users like. Another idea is for the user to tweak their recommendations in the settings. They could prevent certain tags from showing up in recommended. The algorithm should also be tweaked to shy away from the extreme content that users may end up seeing. There will always be a case that YouTube did not account for. These measures are meant to make the algorithm healthier in general.

Acknowledgments I would like to express my thanks to Dr. Pablo Rivas of Marist College who supervised the research and review of this paper. Whenever I needed guidance during the research and writing process, you were there to help. Once again, thank you for your support.

References

1. M. Ledwich, A. Zaitsev, Algorithmic extremism: Examining Youtube's rabbit hole of radicalization, <https://arxiv.org/pdf/1912.11211.pdf>. 19 February, 2020
2. H. Kozłowska, Does YouTube favor radicalization? From outside YouTube, it's hard to know. <https://qz.com/1777381/its-hard-to-know-if-youtubes-algorithm-promotes-radicalization/>. 10 April, 2020
3. "YouTube Regrets," <https://foundation.mozilla.org/en/campaigns/youtube-regrets/>. 28 April, 2020
4. M. Maack, 'YouTube recommendations are toxic,' says dev who worked on the algorithm, <https://thenextweb.com/google/2019/06/14/youtube-recommendations-toxic-algorithm-google-ai/>. 20 February, 2020
5. K. Hao, YouTube is experimenting with ways to make its algorithm even more addictive, <https://www.technologyreview.com/2019/09/27/132829/youtube-algorithm-gets-more-addictive/>. 18 March, 2020
6. P. Cooper, 23 YouTube statistics that matter to Marketers in 2020, <https://blog.hootsuite.com/youtube-stats-marketers/>. 18 March, 2020