

The YouTube algorithm isn't radicalizing people, says bots study

February 20 2024



Credit: Pixabay/CC0 Public Domain

About a quarter of Americans [get their news on YouTube](#). With its billions of users and hours upon hours of content, YouTube is one the largest online media platforms in the world.

In recent years, there has been a popular narrative in the media that videos from highly partisan, conspiracy theory-driven YouTube channels [radicalize young Americans](#) and that YouTube's recommendation algorithm leads users down a path of increasingly radical content.

However, a new study from the Computational Social Science Lab (CSSLab) at the University of Pennsylvania finds that users' own political interests and preferences play the primary role in what they choose to watch. In fact, if the recommendation features have any impact on users' media diets, it is a moderating one.

"On average, relying exclusively on the recommender results in less partisan consumption," says lead author Homa Hosseinmardi, associate research scientist at the CSSLab.

YouTube bots

To determine the true effect of YouTube's recommendation algorithm on what users watch, the researchers created bots that either followed the recommendation engines or completely ignored them. To do this, the researchers created bots trained on the YouTube watch history from a set of 87,988 real-life users collected from October 2021 to December 2022.

Hosseinmardi and co-authors Amir Ghasemian, Miguel Rivera-Lanas, Manoel Horta Ribeiro, Robert West, and Duncan J. Watts aimed to untangle the complex relationship between user preferences and the recommendation algorithm, a relationship that evolves with each video watched.

These bots were assigned individualized YouTube accounts so that their viewing history could be tracked, and the partisanship of what they watched was estimated using the metadata associated with each video.

During two experiments, the bots, each with its own YouTube account, went through a "learning phase"—they watched the same sequence of videos to ensure that they all presented the same preferences to YouTube's algorithm.

Next, bots were placed into groups. Some bots continued to follow the watching history of the real life user it was trained on; others were assigned to be experimental "counterfactual bots"—bots following specific rules designed to separate user behavior from algorithmic influence.

In experiment one, after the learning phases, the control bot continued to watch videos from the user's history, while counterfactual bots deviated from users' real-life behavior and only selected videos from the list of recommended videos without taking the user preferences into account.

Some counterfactual bots always selected the first ("up next") video from the sidebar recommendations; others randomly selected one of the top 30 videos listed in the sidebar recommendations; and others randomly selected a video from the top 15 videos in the homepage recommendations.

The researchers found that the counterfactual bots, on average, consumed less partisan content than the corresponding real user—a result that is stronger for heavier consumers of partisan content.

"This gap corresponds to an intrinsic preference of users for such content relative to what the algorithm recommends," Hosseinmardi says. "The study exhibits similar moderating effects on bots consuming far-left content, or when bots are subscribed to channels on the extreme side of the political partisan spectrum."

'Forgetting time' of recommendation algorithms

In experiment two, researchers aimed to estimate the "forgetting time" of the YouTube recommender.

"Recommendation algorithms have been criticized for continuing to recommend problematic content to previously interested users long after they have lost interest in it themselves," Hosseinmardi says.

During this experiment, researchers calculated the recommender's forgetting time for a user with a long (120 video) history of far-right video consumption who changes their diet to moderate news for the next 60 videos.

While the control bots continued watching a far-right diet for the whole experiment, counterfactual bots simulated a user "switching" from one set of preferences (watching far-right videos) to another (watching moderate videos). As the counterfactual bots changed their media preferences, the researchers tracked the average partisanship of recommended videos in the sidebar and homepage.

"On average, the recommended videos on the sidebar shifted toward moderate content after about 30 videos," Hosseinmardi says, "while homepage recommendations tended to adjust less rapidly, showing homepage recommendations cater more to one's preferences and sidebar recommendations are more related to the nature of the video currently being watched."

"The YouTube recommendation [algorithm](#) has been accused of leading its users toward conspiratorial beliefs. While these accusations hold some merit, we must not overlook that [users](#) have a significant agency over their actions and may have viewed the same content, or worse, even without any recommendations," Hosseinmardi says.

Moving forward, the researchers hope that others can adopt their method

for studying AI-mediated platforms where user preferences and algorithms interact in order to understand better the role that algorithmic content [recommendation](#) engines play in our daily lives.

The findings are [published](#) in the journal *Proceedings of the National Academy of Sciences*.

More information: Homa Hosseinmardi et al, Causally estimating the effect of YouTube's recommender system using counterfactual bots, *Proceedings of the National Academy of Sciences* (2024). [DOI: 10.1073/pnas.2313377121](#)

Provided by University of Pennsylvania

Citation: The YouTube algorithm isn't radicalizing people, says bots study (2024, February 20) retrieved 8 May 2024 from <https://techxplore.com/news/2024-02-youtube-algorithm-isnt-radicalizing-people.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
