YouTube’s Recommendation System

Armin Pousti

Dmitrii Vlasov

Mcgill Department of Computer Science

COMP 480

Professor: Joseph Vybihal

July 24, 2023

In this module, we will delve into an examination of YouTube's recommendation system, exploring its potential impact on promoting radicalizing content and investigating the possibility of a bias favoring right-wing ideologies, which may lead to the emergence of a "rabbit hole" effect toward radicalized right-wing content.In this investigation, we will examine two scientific study pertaining to the subject at hand. The first study, authored by Anna Zaitsev and Mark Ledwich in 2019, constitutes a significant study concerning the YouTube recommendation system. This paper systematically scrutinized four prevailing allegations leveled against the recommendation system: (1) the generation of radical bubbles, (2) the presence of a right-wing bias, (3) the exertion of radicalizing influence, and (4) the facilitation of right-wing radicalization pathways. We shall regard this study as the primary reference, henceforth referred to as the "parent paper." The second scientific study, authored by Mika Desblancs in 2021, is dedicated to a comprehensive analysis of the influence of YouTube algorithms on user siloing. In this the author delves into the effects of YouTube's algorithms on segregating users into distinct and isolated information echo chambers known as silos.

We can dive into how YouTube's recommendation system operates to better understand the topic. This system leverages advanced algorithms and machine learning methodologies to provide personalized content suggestions to individual users. Its principal objective centers on maximizing user engagement by ensuring that the content offered aligns precisely with each user's distinct interests and preferences. Through this tailored approach, YouTube aims to prolong the duration of user interactions on the platform, fostering a mutually beneficial relationship between the user and the platform itself.

The YouTube recommendation system operates on a fundamental principle articulated as "the simple principle of helping people find the video they want to watch and that will give them value" (Goodrow & Cristos, 2021). Over time, this system has evolved significantly, initially relying on popularity-based rankings rather than personalized suggestions. The shift towards personalization stems from YouTube's recognition of the diversity in users' viewing habits, leading the system to analyze and compare these habits with those of similar users to identify potential content that may resonate with each individual. The process by which the YouTube recommendation system work is as follows:

Data Collection:

The initial stage of YouTube's recommendation system involves the comprehensive collection of extensive data pertaining to user behavior. This includes tracking of users' viewing history, search queries, preferences denoted through likes and dislikes, watch time, and various interactions with videos. The system utilizes this vast dataset to construct user profiles, thereby gaining profound insights into individual viewing preferences.

Content Analysis:

Subsequently, YouTube's algorithms undertake an analytical examination of the metadata associated with videos. This analysis includes video titles, descriptions, tags, and the actual video content itself. Additionally, critical factors such as video category and viewer engagement metrics are taken into account, facilitating an objective assessment of each video's quality and relevance within the recommendation framework.

Collaborative Filtering:

As we saw in prior modules, collaborative filtering constitutes a primary technique employed within recommendation systems, and YouTube's recommendation system is no exception to this paradigm. Collaborative filtering techniques are used by the system to identify user similarities based on their distinctive viewing patterns. By discerning patterns of shared interests, the recommendation system proficiently suggests videos that have been favorably received by users exhibiting similar preferences.

Content-Based Filtering:

Within the YouTube recommendation system, content-based filtering is a prominent methodology employed for video suggestions. This technique relies on discerning the inherent attributes and distinct characteristics of videos that a user has previously engaged with, thereby facilitating the generation of tailored recommendations, again as we saw in prior modules.

Engagement Metrics: YouTube effectively utilizes engagement metrics as a vital criterion to ascertain the prominence and resonance of videos. The evaluation encompasses metrics such as likes, comments, shares, and watch time, serving as indicators of positive user feedback. Videos with favorable engagement metrics are prioritized for recommendation, thus augmenting their visibility to a broader audience of potential viewers.

It is essential to note that the system balances exploring new content to introduce users to new topics and exploiting what it knows about users to provide familiar and engaging content. Moreover, the platform also incorporates user feedback and adheres to community guidelines to address concerns related to harmful content and potential biases in the recommendation process.

Ledwich and Zaitsev conducted a scientific study titled "Algorithmic extremism: Examining YouTube's rabbit hole of radicalization." In this study, they examined 816 channels that had more than ten thousand subscribers and focused on channels where at least 30 percent of the content was political. The main objective was to investigate the flow of algorithmic traffic between channels to determine if YouTube's recommendation system promoted radicalized content.

Through their analysis, the researchers categorized the channels and examined the recommendations received by each channel type. The findings of their study contradicted the notion of radicalization, indicating that YouTube's recommendation system actively discourages viewers from engaging with or visiting content that promotes radicalization or extremism.

Furthermore, the data analysis revealed an interesting pattern regarding the algorithm's preferences. The study found that mainstream media and cable news receive preferential treatment over independent YouTube channels, particularly those with a left-leaning bias.

The research then focused on four central claims, with the first claim centering around "radical bubbles." This claim explored how recommendations influence viewers of radical content to consume more content that aligns with their existing beliefs, thereby reducing exposure to alternative viewpoints. The data analysis indicated that the algorithm tends to provide recommendations within the same category or related categories. However, it was observed that this intra-category preference varies based on the channel's category. Notably, channels with potentially radicalizing content exhibited very low recommendation percentages within their own group, indicating a reluctance of the algorithm to reinforce radicalization tendencies within these channels.

The second claim pertained to the "Right Wing Advantage," wherein it was suggested that YouTube's recommendation algorithm exhibited a preference for right-wing content compared to other ideological perspectives. However, the data analysis did not support this claim. Instead, the findings revealed that the recommendation algorithm favored content falling under the mainstream media category.

Notably, the channels experiencing a perceived disadvantage shared a common characteristic: they were operated by independent content creators and not affiliated with broadcasting networks or mainstream journals. This distinction seemed to play a role in their comparatively lower favorability within the algorithm's recommendation system.

The third claim focused on the "YouTube's radicalization influence," suggesting that the algorithm guides users towards more extreme content beyond their usual preferences. However, both the data analysis and the behavior of the recommendation algorithm did not substantiate this claim. On the contrary, the algorithm demonstrated a proactive restriction of traffic towards extreme right-wing categories, particularly channels falling under the classifications of "White Identitarian" and "Conspiracy theory." Notably, these channels received minimal traffic, indicating a deliberate effort by the algorithm to limit their visibility and dissemination.

The final claim examined the "Right-wing Radicalization pathway" on YouTube, proposing that the algorithm guides viewers of mainstream and center-left channels towards progressively more left-wing critical content, eventually leading them to extreme right-wing content. However, the empirical data presented contrary evidence to support this claim. The data also indicated that the only right-wing media entity benefiting the most from the algorithm was Fox News, a prominent mainstream right-wing outlet. The preferential treatment of Fox News was attributed to the relative scarcity of other right-leaning mainstream media sources. On the other hand, traffic concerning Center/Left mainstream media was found to be more evenly distributed across representative outlets within that category. This finding implies a differentiated impact of the algorithm on media engagement patterns based on the availability of mainstream outlets in the respective ideological domains.

Desblancs conducted a second study focusing on the siloing effects of an automated driver following a recommendation system to examine potential right-wing bias and radicalization. The study employed a scrapper that started with a seed video, either from a partisan right or partisan left perspective, and then proceeded to follow YouTube's recommendations. The scrapper visited five layers in the tree structure of recommended videos, taking three recommendations for each video. Metadata from each video was collected to categorize them as either right-wing, center, or left-wing in terms of their political affiliation. The findings indicated that the recommendation system exhibited a higher frequency of favoring left-wing videos.

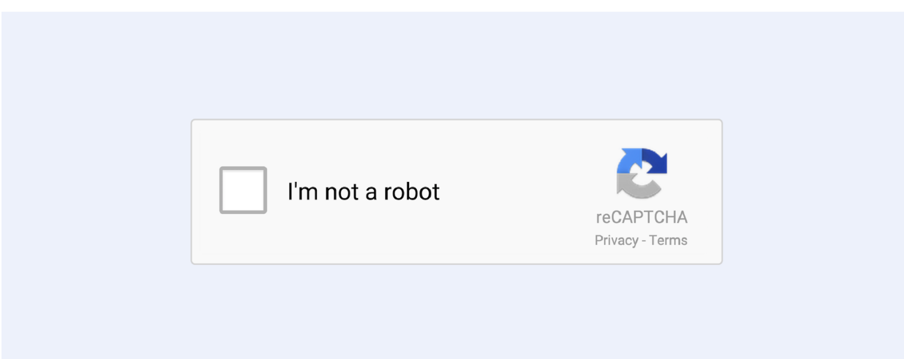
Reimplementation

As a result of evolving web design and Google's ongoing battle against bots, numerous methods and functions in the original code have become inoperable. To address this challenge and gain a deeper comprehension of the code's functionality, essential parts were implemented from the ground up.

This, however, involved solving several issues:

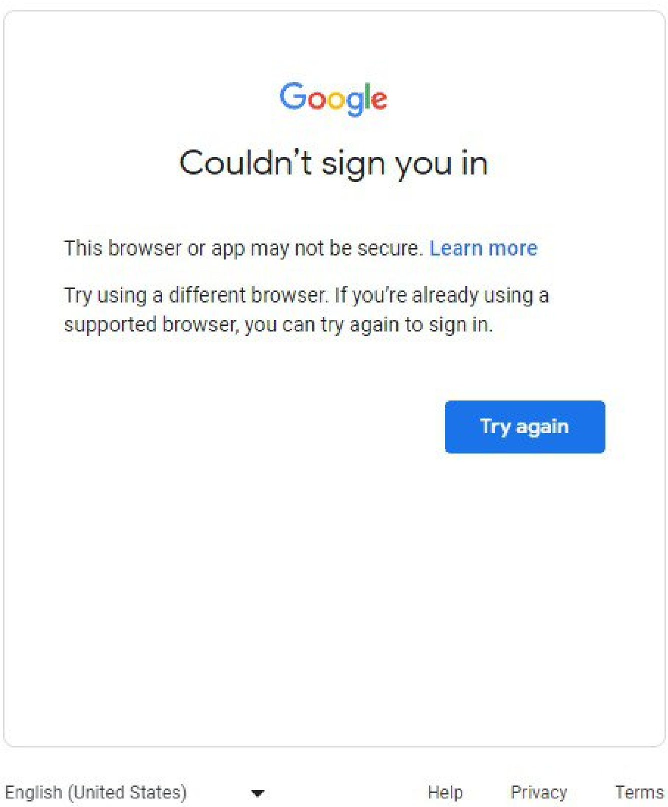
1) *Bot flag*: In order to overcome being flagged as a bot, the attached implementation introduced explicit wait commands in between clicks in order to better mimic human behavior. Strangely enough, sending email and password characters all at once or one at a time produced the same result.

2) *Captcha*: Several attempts of solving captchas were tried. Since the goal was full automation, stopping the code for manual captcha solution was not considered. One of the ways to deal with this issue was to use AutoClicker programs to solve "I am not a robot" captchas:



Since this is not the only type of captcha, whenever a click would not result in a successful login to YouTube, the code simply restarted to try to get this "easy" captcha.

3) *App may not be secure:* Even after bypassing captchas, it seems that Google is now capable of detecting the use of Selenium or ChromeDriver. This proved to be the greatest obstacle since it had nothing to do with human/bot behavior, but rather the library used. Neither of the ChromeOptions attributes seemed to have any effect on this behavior. The solution to this problem lies in the use of the default chrome driver. Instead, the reimplementation uses undetected-chromedriver 3.5.0 whose whole purpose is to not trigger anti-bot services like Distill Network / Imperva / DataDome / Botprotect.io. The use of this library resolved all issues and did not trigger captchas. Therefore the aforementioned AutoClicker solution to captchas became obsolete.



4) *Xpaths and Java Script:* Since YouTube design changed, the xpaths needed to be modified. However, certain buttons (like the History button) could not be clicked using xpaths, as mentioned above, therefore JS syntax was used instead.

Functionality

1. Log-in
2. History Deletion
3. Watching the seed video (potentially a list of seed videos)
4. DFS following the recommendation tree. Starting from a seed video, watch the first recommended video for each video, while storing top n recommendations. When the final depth of the tree is reached, the algorithm returns to the previous node (DFS) until all children videos are watched up to a specific depth.
5. Currently, each video’s recommendation is printed, however, the code could be easily modified to store the results in a tree.

This code foundation opens the possibility to recreate the original code, while solving many issues.

Work Cited

- “View of Algorithmic Extremism: Examining YouTube’s Rabbit Hole of Radicalization: First Monday.” *View of Algorithmic Extremism: Examining YouTube’s Rabbit Hole of Radicalization | First Monday*, firstmonday.org/ojs/index.php/fm/article/view/10419/9404. Accessed 7 Aug. 2023.
- Mika-Jpd. “Mika-Jpd/Youtube_radicalization_recommendations: Chapter 4: Analysis of the Impact of Algorithms in Siloing Users: Special Focus on YouTube for the Book ‘Ai & Society: Tensions and Opportunities.’” *GitHub*, github.com/mika-jpd/YouTube_Radicalization_Recommendations/tree/master. Accessed 7 Aug. 2023.
- Goodrow, Cristos. “On YouTube’s Recommendation System.” *Blog.Youtube*, 15 Sep. 2021, blog.youtube/inside-youtube/on-youtubes-recommendation-system/#:~:text=Recommendations%20on%20%22Up%20next.%22,you%20may%20want%20to%20watch.
- Kumar, P. Pavan, et al. *Recommender Systems: Algorithms and Applications*. CRC Press, 2021.
- Robillard, Martin P., et al. An Introduction to Recommendation Systems in Software Engineering. Springer Science and Business, 2014.