SURFING YOUTUBE: LOOKING FOR ALGORITHMIC OVERDEPENDENCE USING SIMULATED DIGITAL CONTENT PLATFORM RECOMMENDERS

Walter Stover

GEORGE MASON UNIVERSITY

**I Introduction**

In the digital age, recommender systems play a pivotal role in guiding users through vast amounts of choices and content, especially on digital media platforms such as YouTube. While recommender systems aim to enhance user experience by suggesting relevant content, there has been experimental evidence suggesting that users can become overly dependent on these systems and following recommendations when they shouldn’t. This has been explored primarily in traditional consumer platforms such as Amazon, and this paper explores whether this phenomenon of algorithmic overdependence can also occur on digital media platforms by framing the search for optimal video content as a Weitzman (1979) “Pandora’s Box” optimal search problem and embedding this problem in an agent-based model (ABM) to simulate users watching videos.

**II Background**

Earlier studies have shown mixed results for the effect of recommender systems on users – these systems can enhance decisions by reducing search effort, or lead to users being worse off if they follow recommendations when they would have been better off using their own judgment. This paper specifically extends the work of Banker and Khetani (2019), which provided experimental evidence of users' decision quality declining due to excessive trust in recommender systems, which they termed *algorithmic overdependence*.

However, Banker and Khetani conducted their experiment in a traditional consumer environment where consumers choose between different products between which strict technical dominance could be established. Their experiment asked subjects to make a purchase decision from a set of portable chargers, with and without automated product recommendations, and found that consumers on average were susceptible to following inferior recommendations. The goal of this paper is to investigate whether this algorithmic overdependence can also occur on a digital media platform such as YouTube, where users choose between videos where strict dominance between options cannot be established – the value of any particular video is subjective.

**III Methodology**

This paper introduces an agent-based model that simulates platform users navigating and watching videos. This simulation mirrors a randomized trial with a control group of virtual users who independently navigate the platform in search of optimal video content, and two treatment groups faced with the same problem, but with a virtual recommender system providing suggestions. In the first treatment, the recommender system suggests videos at random, while in the second treatment, the recommender system calculates and recommends the highest-value videos within the users' search vicinity.

I model the problem of finding the best video content as an optimal search problem following Weitzman (1979)’s “Pandora’s Box” framework. Videos then have some value and cost of watching that are drawn from a known cumulative distribution function. Users attempt to maximize the expected total value of the videos’ content, minus the costs of watching them. In the ABM, these videos are distributed across a 20x20 cell grid that users navigate. Each cell contains a video that the user will then watch. Users are aware of videos in their immediate 8-cell Moore neighborhood and form expectations of the value of these neighboring videos with some accuracy according to a soft-max calculation.

In the ABM, users are initialized with heterogeneous values of *acuity*, *recommender trust*, and *patience*. *Acuity* determines the ability of the user to calculate the expected value of the videos in their neighborhood; *recommender trust* determines how likely the user is to select the option suggested by the recommender; finally, *patience* determines how tolerant the user is of payoff losses before they exit the search.

Running the agent-based model consists of one session per virtual user, with multiple rounds per session. One session generates the distribution of videos across the grid and places one virtual user at a random location. Each round, the user will:

1. Calculate whether to stop searching.

2. Move to a new cell in their neighborhood according to their soft-max calculation of which cell contains the highest value video. Soft-max calculation takes in expected value scores of a consideration set, and exponentiates each score before normalizing it to sum to 1, such that the options with higher scores are given exponentially more weight in the resulting probability distribution.

3. Watch the video and record the resulting payoff in their memory.

4. Calculate average payoff according to their current payoff memory.

5. Calculate payoff direction of whether this video had a higher or lower payoff than the previous video they watched.

In the two treatments, users will also choose whether to follow the recommenders’ choice instead of their own.

**IV Key Findings**

The random recommender treatment was statistically significant in their impact on user decisions at the 10% level, but with a small negative magnitude. The high-valuation recommender treatment had a similarly small, negative magnitude, but did not obtain statistical significance. How long an agent is willing to search – their *patience* factor – is statistically significant at the 1% level, and has a much stronger association with greater final payoffs than the accuracy of their judgment of the alternatives as represented by *acuity*, or the type of recommender system in place.

Thus, any type of recommender system makes them worse off on average, even the high-valuation system that knows the maximum value video to recommend to users. *Recommender trust* is significant at the 10% level and negative, with a small coefficient. Comparing this to Banker and Khetani (2019), these results do show limited evidence for algorithmic overdependence, as recommender systems tend to lead to worse payoffs for users overall. A negative sign for *recommender trust* is also consistent with findings from their treatment comparing different levels of trust in a recommender system.

However, my results differ in three main respects. First, overall, the effect of recommender systems on decision quality is not as significant as in Banker and Khetani’s findings – instead, patience is the most significant and impactful factor. Second, even high-valuation recommenders lead to worse user payoffs on average, which was not one of Banker and Khetani’s treatments. In the original study, algorithmic overdependence occurred because the recommender system deliberately recommended technically dominated options, and users committed type I errors by adopting recommendations that made them worse off.

This suggests that it’s possible that algorithmic overdependence can occur even if the recommender system knows and recommends the highest-valued option in a user’s immediate neighborhood. However, this high-valuation recommender treatment did not obtain statistical significance and therefore, I cannot conclude that this form of algorithmic overdependence is actually occurring. It is possible that this treatment did not obtain significance because both users and recommenders are limited to calculating value of the options in the users’ immediate neighborhood, instead of being allowed to consider options across the entire grid.

**V Conclusion**

In this paper, I used an agent-based modeling approach combined with a Weitzman-type Pandora’s Box model to simulate viewer consumption of media on YouTube and other digital media platforms as an optimal search problem. The results show some limited evidence for algorithmic overdependence, but patience is a much more important factor. This may be due to restricting users in the ABM to searching in their immediate neighborhood as opposed to searching across the entire grid.

Further steps to build on this paper’s results would include revising the ABM to allow users to move anywhere on the grid at will, allowing evaluation of all alternatives and widening the consideration set of each watcher. This would allow me to more robustly test for algorithmic overdependence, and test if the high-valuation recommender treatment can obtain significance. Additionally, in this revised version, I would engage in more thorough hyper-parameter space search to test models for each possible combination of parameter values for users. Finally, I want to compare these findings with data from a real-world human subject experiment with the same control and treatment groups.

**References**

Weitzman, Martin L. “Optimal Search for the Best Alternative.” *Econometrica* 47, no. 3 (1979): 641–54.

Banker, Sachin, and Salil Khetani. “Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being.” *Journal of Public Policy & Marketing* 38, no. 4 (October 1, 2019): 500–515.