INFLUENCIANDO RECOMENDAÇÕES DO YOUTUBE POR MEIO DE LINKS COMPARTILHADOS

Adriano Cabral Linhares Mourthé [1]; Carlos E. Mello [1]

1 - Filiação –Programa de Pós-Graduação em Informática, Universidade Federal do Estado do Rio de Janeiro (UNIRIO).

Apoio Financeiro: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

RESUMO: Este trabalho investiga o impacto das interações com links externos, especificamente aqueles compartilhados via WhatsApp, nas recomendações do YouTube. Foi desenvolvida uma metodologia baseada em simulação, utilizando perfis automatizados para replicar o comportamento de usuários reais. Esses perfis foram expostos a links externos, e as mudanças nas recomendações do YouTube resultantes dessas interações foram analisadas para avaliar o impacto dos links no algoritmo de recomendação.

Os resultados mostram que a exposição a esses links pode alterar as recomendações oferecidas pelo YouTube, sugerindo um potencial de manipulação dos conteúdos recomendados.

Palavras-chave: Sistemas de recomendação; plataformas digitais; manipulação de algoritmos; *sock puppets.*

INTRODUÇÃO: Plataformas digitais como YouTube, Instagram e TikTok consolidaram-se como os principais espaços para a criação e distribuição de conteúdo. No centro dessas plataformas, os Sistemas de Recomendação (SRs), baseados em algoritmos, atuam como a principal tecnologia de personalização de conteúdo para os usuários. Esse processo de curadoria algorítmica não apenas seleciona conteúdos relevantes, mas também influencia diretamente a formação de opiniões (JANNACH; JUGOVAC, 2019).

De acordo com pesquisa publicada pelo Instituto de Pesquisa Data Senado, 40% dos brasileiros já decidiram seu voto com base em informações obtidas em redes sociais, como Facebook, Twitter, Instagram, WhatsApp ou YouTube. O WhatsApp, em particular, destaca-se como uma das principais plataformas para a disseminação de notícias e compartilhamento de links (BRASIL, 2019), incluindo vídeos do YouTube.

Este trabalho investiga o impacto do compartilhamento de links externos, especificamente via WhatsApp, nas recomendações personalizadas oferecidas pelo YouTube. Para abordar essa questão, desenvolvemos uma metodologia baseada em simulações utilizando perfis automatizados, conhecidos como sock puppets, que replicam o comportamento de usuários reais (SANDVIG et al., 2014). Esses perfis simulados permitem a coleta controlada de dados, fornecendo uma base empírica para avaliar como as recomendações do YouTube são influenciadas por interações externas.

OBJETIVOS: Este trabalho tem como objetivo desenvolver uma metodologia para investigar o impacto de interações com links externos nas recomendações de usuários no YouTube. Além disso, investigaremos tais efeitos no domínio de esporte e conteúdo infantil.

METODOLOGIA: Foi desenvolvido um ambiente de simulação utilizando sock puppets—bots automatizados que replicam o comportamento de usuários reais no YouTube (SANDVIG et al., 2014). Esses sock puppets foram programados para realizar ações predefinidas, como assistir a vídeos, clicar em links e inscrever-se em canais, simulando a interação típica de um usuário. Nesse ambiente, foram coletados todos os dados relacionados às listas de recomendação oferecidas aos usuários simulados.

Com o objetivo de identificar e mensurar o efeito dos *links* externos nas recomendações do YouTube, estabelecemos dois grupos de *sock puppets*: um grupo de controle, que seguiu um padrão de navegação sem exposição a links externos, e um grupo de tratamento, que foi intencionalmente exposto a esses links. A comparação entre as recomendações geradas para cada grupo permitiu estimar o impacto dos links externos nas sugestões do YouTube, utilizando uma abordagem análoga à estimativa do efeito de um tratamento (CUNNINGHAM, 2021).

O experimento de simulação foi conduzido em dois passos. No primeiro passo, ambos os grupos, controle e tratamento, seguiram o mesmo padrão inicial de consumo, replicando o comportamento típico de navegação no YouTube. No segundo passo, os grupos foram tratados de forma diferente: o grupo de controle manteve seu perfil de consumo esperado, enquanto o grupo de tratamento foi exposto a links externos que direcionavam para um vídeo de cunho político no YouTube.

Os links externos utilizados para o grupo de tratamento foram selecionados com base em dados extraídos do estudo de Marres e Welbers (2019), que analisou mensagens de grupos partidários no WhatsApp durante a eleição presidencial brasileira de 2018. A escolha desses dados se deu por sua relevância política e pelo potencial de enviesar as listas de recomendação, direcionando-as para o conteúdo do vídeo do link.

Um aspecto adicional avaliado nas recomendações foi a similaridade entre as listas de recomendação do grupo de controle e as do grupo de tratamento. No contexto do experimento, essa similaridade indica se há ou não uma divergência entre as listas de recomendação dos dois grupos.

No entanto, avaliar a similaridade entre listas de recomendações não é uma tarefa simples, uma vez que essas listas são compostas por nomes de vídeos e canais escritos em linguagem natural. Para lidar com esse desafio, todos os canais recomendados para os usuários foram integrados em uma matriz de utilidade, onde cada dimensão representa o número de vezes que um canal aparece em uma lista de recomendações. Com as listas de recomendação representadas de forma vetorial, é possível empregar a similaridade cosseno para estimar a similaridade entre as diferentes listas (LUDEWIG; JANNACH, 2018).

RESULTADOS: A análise dos dados coletados revelou que a exposição a um vídeo de cunho político produziu mudanças significativas nas recomendações do YouTube, tanto em termos de conteúdo quanto de canais sugeridos. Essas alterações foram observadas em ambos os perfis analisados, mas com magnitudes distintas.

No perfil infantil, antes da intervenção, 66% dos vídeos e 81% dos canais recomendados eram comuns entre os grupos de controle e tratamento. Esse alto nível de interseção indica que, inicialmente, as recomendações eram bastante consistentes entre os grupos. No entanto, após a exposição ao vídeo político, essa consistência diminuiu drasticamente. Especificamente, a interseção nas recomendações de vídeos foi reduzida em 93,4%, enquanto a interseção nas recomendações de canais caiu 86,9%. Esses dados sugerem que o algoritmo de recomendação do YouTube passou a priorizar conteúdos diferentes para os dois grupos, influenciado fortemente pelo vídeo político.

Além disso, observou-se que a média de vídeos recomendados do canal de intervenção para o perfil infantil foi de 2,54. Isso evidencia que, além de reduzir a interseção das recomendações, o algoritmo também passou a sugerir mais frequentemente conteúdos relacionados ao canal do vídeo político, indicando uma possível formação de bolhas de conteúdo ao redor desse tema.

Para o perfil de esportes, a intervenção com o vídeo político também resultou em um efeito, embora menos pronunciado em comparação ao perfil infantil. Inicialmente, 52% dos vídeos e 53% dos canais recomendados eram comuns entre os grupos de controle e tratamento, refletindo uma certa consistência nas recomendações. Após a intervenção, essa interseção diminuiu significativamente, com uma redução de 87,7% na interseção de vídeos e de 81% na de canais. Assim como no perfil infantil, a exposição ao vídeo político desviou consideravelmente as recomendações do algoritmo.

A média de vídeos do canal de intervenção recomendados para o perfil de esportes foi de 2,19. Embora menor do que no perfil infantil, esse número ainda indica uma alteração nas recomendações. Esses resultados reforçam a ideia de que, mesmo em perfis com interesses bem definidos, como o de esportes, a exposição a conteúdos politicamente carregados pode modificar as sugestões oferecidas pelo algoritmo do YouTube.

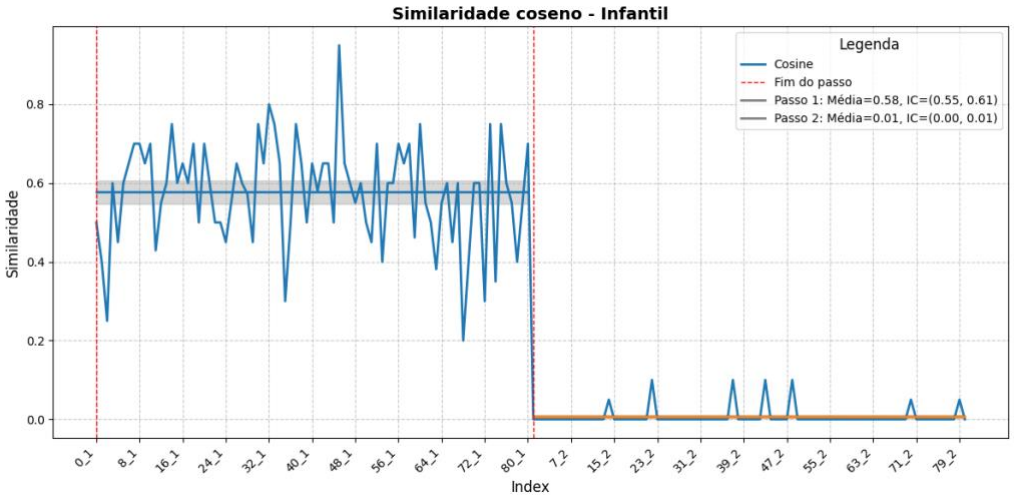

Figura 1 - Gráfico da similaridade cosseno para os perfis infantil
Fonte: Elaboração própria.

Em relação à similaridade entre as listas de recomendações dos grupos de tratamento e controle, como esperado, ao perturbar a entrada do sistema de recomendação (SR) usando o link externo, o SR respondeu alterando a distribuição das recomendações. Isso é indicado pela diminuição da similaridade observada nas figuras 1 e 2, que mostram a comparação entre os grupos de controle e tratamento para os perfis do experimento.

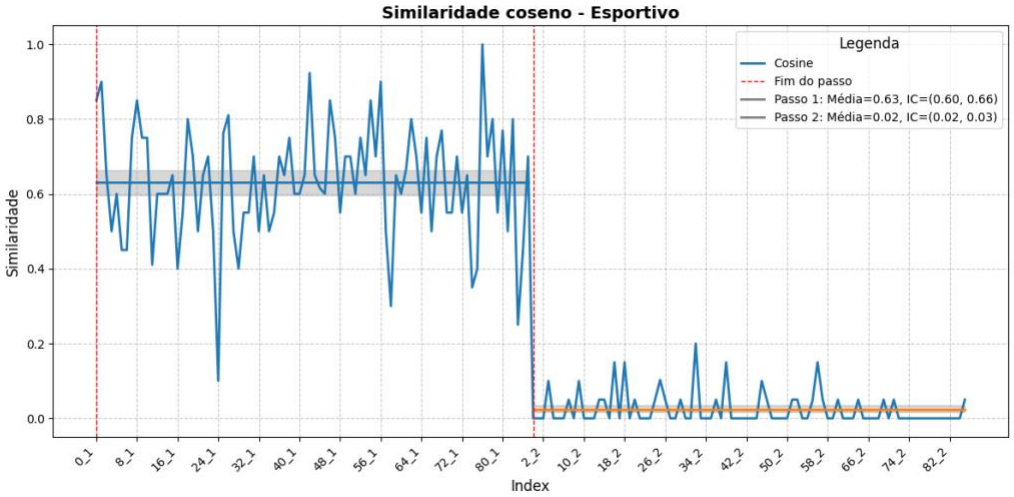

Figura 2 - Gráfico da similaridade cosseno para os perfis esportivos
Fonte: Elaboração própria.

No perfil infantil, a similaridade cosseno entre os passos da simulação apresentou uma variação de 0,58 antes da intervenção e 0,01 após a exposição ao vídeo de cunho político. De maneira semelhante, no perfil esportivo, a similaridade cosseno variou de 0,63 para 0,02. Esses resultados indicam o impacto que o *link externo* teve na percepção do sistema de recomendação sobre as preferências dos usuários.

CONCLUSÕES: Este trabalho investigou o impacto de links externos nas recomendações do YouTube por meio de uma metodologia baseada em simulações com sock puppets. Os resultados forneceram evidências quantitativas de que a exposição a links externos pode alterar o perfil de recomendações oferecido aos usuários, tanto em termos de conteúdo sugerido. Essa influência, observada em diferentes domínios, como esportes e conteúdo infantil.

Os resultados obtidos mostram como a exposição a um único vídeo político pode impactar as recomendações feitas pelo YouTube. Essa divergência foi mais pronunciada no perfil infantil, mas também foi evidente no perfil de esportes. Esses achados são importantes para a compreensão de como os mecanismos de recomendação podem ser influenciados por conteúdos externos ao sistema, com possíveis implicações para a formação de bolhas de conteúdo e a amplificação de conteúdo político.

Futuros trabalhos podem aprofundar a investigação sobre o impacto de links externos em outros domínios e plataformas, além de explorar estratégias para mitigar seus efeitos negativos e promover uma experiência online mais diversa e segura para os usuários.

REFERÊNCIAS BIBLIOGRÁFICAS:

JANNACH, Dietmar; JUGOVAC, Michael. Measuring the business value of recommender systems. ACM Transactions on Management Information Systems (TMIS), v. 10, n. 4, p. 1-23, 2019.

BRASIL. Senado Federal. Instituto de Pesquisa DataSenado. Acesso à internet e comportamento nas

redes sociais: notícias falsas, privacidade de dados, e perfil dos respondentes. Brasília: Senado Federal, 2019. 113 p. Disponível em: https://www2.senado.gov.br/bdsf/handle/id/603465. Acesso em: 21 ago. 2024.

SANDVIG, Christian; et al. Auditing algorithms: Research methods for detecting discrimination on internet platforms. Data and discrimination: converting critical concerns into productive inquiry, v. 22, p. 4349-4357, 2014.

CUNNINGHAM, Scott. Causal inference: The mixtape. New Haven: Yale University Press, 2021.

MARRES, N.; WELBERS, K. Thousands of Small, Constant Rallies: A Large-Scale Analysis of Partisan WhatsApp Groups. Proceedings of the Intxernational AAAI Conference on Web and Social Media, v. 13, p. 243-252, 2019

GOOGLE. Google Trends. Disponível em: https://trends.google.com. Acesso em: 22 ago. 2024.

LUDEWIG, M.; JANNACH, D. Evaluation of session-based recommendation algorithms. User Modeling and User-Adapted Interaction, v. 28, p. 331-390, 2018.