



## PSICOLOGIA SOCIAL E FAKE NEWS: MODELOS DE IA PARA A PREVENÇÃO DA MANIPULAÇÃO DE MASSAS

Joaquim Antonio de Bastos Filho  
Bruno Slusarz  
Bruno Henrique Guzzo  
Winicius Mainardes  
Taline lenk

**Resumo:** No cerne de inúmeros dilemas digitais contemporâneos e, conseqüentemente, da desinformação, está a relação entre psicologia social, fake news e inteligência artificial (IA). Mais especificamente, a disseminação de fake news por meio de redes sociais e plataformas digitais. As maneiras pelas quais os algoritmos servem como condutores para a propagação de notícias falsas são um dos principais tópicos abordados neste artigo. Esses algoritmos, que visam aumentar o engajamento, acabam por consolidar ainda mais os vieses cognitivos e sociais, estimulando a disseminação de desinformação de maneira mais eficaz e veloz (Pariser, 2011). As descobertas ressaltam que esses algoritmos exploram vulnerabilidades psicológicas, como o viés de confirmação, o que reforça a necessidade de uma abordagem interdisciplinar para combater a manipulação em massa. A metodologia do estudo é baseada em uma revisão de literatura, proporcionando uma visão teórica abrangente do problema. A análise de trabalhos sobre IA e psicologia social permitiu identificar os mecanismos pelos quais as fake news se espalham e como a IA pode contribuir, tanto para ampliar quanto para combater essa disseminação. No entanto, a metodologia é limitada pela ausência de dados empíricos ou experimentos que validem as hipóteses levantadas. Isso impede uma análise mais profunda sobre os impactos reais dos modelos de IA na prevenção de desinformação (Shu et al., 2017). Uma das principais descobertas da pesquisa é que os algoritmos de IA criam bolhas informacionais, onde os usuários são expostos apenas a conteúdos que reforçam suas próprias crenças. Esse processo exacerbado pelo viés de confirmação, dificulta a troca de ideias entre diferentes grupos e fomenta a polarização social (Pennycook & Rand, 2018). A revisão também destaca que as fake news tendem a se espalhar mais rapidamente que as notícias verdadeiras devido ao seu apelo emocional, tornando ainda mais complexo o combate à desinformação (Vosoughi, Roy & Aral, 2018). Além disso, o estudo considera o efeito backfire, que ocorre quando a tentativa de corrigir mitos ou desinformação leva, paradoxalmente, ao fortalecimento das crenças equivocadas. Esse fenômeno representa um grande obstáculo na luta contra a desinformação, especialmente em contextos polarizados (Nyhan & Reifler, 2010). O estudo sugere que, para superar esse desafio, os modelos de IA precisam ser ajustados para personalizar estratégias de correção, levando em consideração as vulnerabilidades psicológicas e sociais dos usuários. No entanto, essa abordagem enfrenta desafios éticos, como a privacidade dos dados pessoais e o uso de perfis comportamentais. Em conclusão, embora a IA tenha grande potencial para combater a disseminação de fake news, sua eficácia depende da incorporação de insights psicológicos em seus algoritmos. O estudo ressalta que uma colaboração interdisciplinar entre psicologia social e IA é essencial para enfrentar os desafios éticos e práticos na

prevenção da manipulação de massas (Sunstein, 2009). A abordagem proposta, ao mesmo tempo promissora, exige um maior desenvolvimento empírico para validar suas aplicações e impactos na sociedade.

**Palavras-chave:** Inteligência Artificial. Fake News. Psicologia Social.

Referências:

Nyhan, B., & Reifler, J. (2010). **Quando as correções falham: A persistência de percepções políticas equivocadas**. *Comportamento Político*, 32(2), 303–330.

Pariser, E. (2011). **O filtro invisível: O que a Internet está escondendo de você**. Penguin Press.

Pennycook, G., & Rand, D. G. (2018). **O efeito de verdade implícita: Atribuir avisos a um subconjunto de histórias falsas aumenta a precisão percebida de histórias sem avisos**. *Ciência da Administração*, 66(11), 4944–4957.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). **Deteção de notícias falsas em mídias sociais: Uma perspectiva de mineração de dados**. *Boletim de Explorações da ACM SIGKDD*, 19(1), 22–36.

Sunstein, C. R. (2009). **Sobre boatos: Como as falsidades se espalham, por que acreditamos nelas e o que pode ser feito**. Princeton University Press.

Vosoughi, S., Roy, D., & Aral, S. (2018). **A disseminação de notícias verdadeiras e falsas online**. *Ciência*, 359(6380), 1146–1151.