

Instâncias textuais do racismo algorítmico

Por Júlio Araújo

Fonte: OpenAlex · UFC · [Publicação original](#)

INOVAÇÃO	APLICABILIDADE	POTENCIAL ECONÔMICO	MATURIDADE
8/10	9/10	8/10	Média

Este paper analisa algoritmos de IA como textos estruturantes que reproduzem estereótipos raciais, sustentando dinâmicas de exclusão. Propõe um modelo teórico (Prompt-Text, Response-Text, Algorithm-Text) e demonstra, via simulação em Python, como algoritmos associam termos racializados à violência. Defende a reescrita algorítmica como um ato ético e político, propondo estratégias como diversificação de datasets, filtros de viés, transparência e regulação para dismantelar lógicas excludentes.

NÚCLEO DA ANÁLISE

Ideia de startup ou produto

Uma startup focada em 'Auditoria e Reescrita Algorítmica'. Ofereceria uma plataforma SaaS para empresas avaliarem e corrigirem vieses em seus modelos de IA, especialmente em geração de conteúdo (texto, imagem). A plataforma incluiria ferramentas de análise de datasets, filtros de viés pré-treinados e módulos para 'reescrever' comportamentos algorítmicos indesejados, garantindo outputs mais equitativos e representativos.

Aplicações práticas

Desenvolvimento de ferramentas de auditoria e mitigação de viés para sistemas de IA, criação de diretrizes para o design ético de algoritmos, implementação de filtros de preconceito em modelos de IA generativa, e formação de equipes multidisciplinares para o desenvolvimento de IA mais justa e equitativa.

Potencial de mercado

O mercado para soluções de IA ética e mitigação de viés é crescente, impulsionado por regulamentações emergentes (e.g., AI Act da UE), demanda por responsabilidade corporativa e a necessidade de evitar danos reputacionais. Empresas de tecnologia, agências governamentais e setores regulados são potenciais clientes para ferramentas e consultoria em IA justa e transparente.

FUNDAMENTO CIENTÍFICO

Problema abordado

O problema central abordado é a reprodução sistemática de racismo e estereótipos raciais por algoritmos de inteligência artificial, resultando em dinâmicas de exclusão e reforço de desigualdades históricas, muitas vezes mascarado pela atribuição de responsabilidade ao usuário ou aos dados, e não ao algoritmo em si.

Metodologia

O autor desenvolve um modelo teórico de três instâncias textuais (Prompt-Text, Response-Text, Algorithm-Text), analisa um caso de viés em imagem gerada por IA, e utiliza uma simulação em Python para demonstrar a associação algorítmica de termos racializados com violência. Complementa com análise de comentários em redes sociais e propõe estratégias de reescrita algorítmica.

Principais descobertas

Algoritmos de IA operam como agentes discursivos que reforçam desigualdades históricas; a associação de termos racializados com violência é sistemática; a responsabilidade pelo viés é frequentemente desviada do algoritmo para o usuário; a reescrita algorítmica (diversificação de datasets, filtros de viés, transparência, regulação) é essencial para dismantelar lógicas excludentes.

LEITURA PARA GESTÃO PÚBLICA

O estudo pode subsidiar a criação de políticas públicas e marcos regulatórios para IA no Brasil, focando em transparência algorítmica, auditoria obrigatória de sistemas de IA de alto risco, diretrizes para diversificação de datasets e a implementação de mecanismos de responsabilização para empresas que desenvolvem e utilizam IA. Poderia influenciar a criação de um 'selo de IA ética' ou 'conformidade algorítmica'.

PARCERIA UNIVERSIDADE × EMPRESA

Parceria entre a UFC (especialmente departamentos de Ciência da Computação, Direito e Comunicação) e empresas de tecnologia (e.g., desenvolvedores de grandes modelos de linguagem, plataformas de mídia social, empresas de consultoria em IA) para pesquisa aplicada em detecção e mitigação de viés algorítmico. O objetivo seria desenvolver e testar em ambiente real as estratégias de 'reescrita algorítmica' propostas, criando soluções inovadoras e publicações conjuntas.

OPORTUNIDADES DERIVADAS

4 direções estratégicas identificadas

STARTUP · IMPACTO ALTO · INTELIGÊNCIA ARTIFICIAL

Plataforma de Auditoria e Mitigação de Viés Algorítmico

Desenvolvimento de uma plataforma SaaS para identificar, analisar e corrigir vieses em algoritmos de IA, com foco em modelos generativos de texto e imagem, oferecendo ferramentas para diversificação de datasets e aplicação de filtros de preconceito.

PARCERIA · IMPACTO ALTO · INTELIGÊNCIA ARTIFICIAL

Centro de Pesquisa Aplicada em IA Ética e Inclusiva (UFC-Indústria)

Colaboração entre a UFC e empresas de tecnologia para desenvolver metodologias e ferramentas para o design, desenvolvimento e avaliação de sistemas de IA que sejam intrinsecamente éticos, transparentes e livres de vieses, com foco em casos de uso reais.

POLÍTICA PÚBLICA · IMPACTO ALTO · GOVTECH

Marco Regulatório para Transparência e Responsabilidade Algorítmica

Elaboração de propostas de políticas públicas e regulamentações que exijam auditorias de viés para sistemas de IA, promovam a transparência algorítmica e estabeleçam mecanismos de responsabilização para a reprodução de estereótipos e discriminação por IA.

PRODUTO CORPORATIVO · IMPACTO MÉDIO · SOFTWARE

Módulos de 'Reescrita Algorítmica' para Ferramentas de Desenvolvimento de IA

Integração de funcionalidades de detecção e correção de viés diretamente em SDKs e APIs de plataformas de desenvolvimento de IA, permitindo que desenvolvedores incorporem 'reescrita algorítmica' como parte do ciclo de vida de seus produtos.

ABSTRACT ORIGINAL

In this article, I examine artificial intelligence (AI) algorithms as structuring texts that process information, shape narratives, and reproduce racial stereotypes, thereby sustaining dynamics of exclusion. I present a theoretical model composed of three interconnected textual instances: the Prompt-Text, which expresses the user's intent; the Response-Text, which materializes the generated output; and the Algorithm-Text, the invisible layer that defines the system's internal rules. I analyze a case in which an AI-generated image depicted a Black woman holding a gun in a favela, despite the original prompt specifying a positive figure. To explore this bias, I developed a Python simulation that demonstrates how the algorithm automatically associates racialized terms with violence, revealing the systematic reproduction of stereotypes. I also examine social media comments that naturalize such distortions, shifting responsibility to the user while obscuring the algorithm's central role in constructing representations. I argue that these systems operate as discursive agents that reinforce historical inequalities and hierarchies. I defend algorithmic rewriting as an ethical and political act, supported by strategies such as diversifying datasets, applying bias filters, ensuring greater transparency, and promoting regulation. I conclude by reaffirming the thesis that guides this work: the algorithm is a text, and, as such, it can – and must – be rewritten to dismantle exclusionary logics.

PALAVRAS-CHAVE

Rewriting · Structuring · Politics · Inequality · Object (grammar) · Process (computing)