

TREINAMENTO DE INTELIGÊNCIA ARTIFICIAL COM DADOS DE NAVEGAÇÃO EM REDES SOCIAIS: ABORDAGENS, IMPLICAÇÕES ÉTICAS E IMPACTO NA PRIVACIDADE DOS USUÁRIOS

ARTIFICIAL INTELLIGENCE TRAINING WITH SOCIAL MEDIA NAVIGATION DATA: APPROACHES, ETHICAL IMPLICATIONS AND IMPACT ON USER PRIVACY

Claudia Silva de Oliveira Camargo¹

Resumo: Com o crescimento das redes sociais, as plataformas digitais adquiriram um volume inédito de dados gerados pelos usuários, incluindo textos, imagens, vídeos e interações sociais. Estes dados são frequentemente utilizados para o treinamento de modelos de Inteligência Artificial (IA), que podem aprender a partir de linguagem, comportamento e preferências humanas. Embora essa prática possibilite avanços significativos na personalização de conteúdo e moderação de redes, ela também levanta questões éticas, especialmente em relação à privacidade dos usuários e ao uso adequado de seus dados pessoais. Este artigo examina os métodos de coleta e tratamento de dados de redes sociais, os principais modelos e algoritmos aplicados e as implicações éticas associadas ao uso dessas informações. Conclui-se com uma análise sobre o impacto dessa prática na privacidade dos usuários e recomendações para uma abordagem mais ética e transparente.

Palavras-Chave: inteligência artificial; treinamento de IA; privacidade; redes sociais; dados de navegação.

Abstract: With the growth of social networks, digital platforms have acquired an unprecedented volume of data generated by users, including texts, images, videos and social interactions. This data is often used to train Artificial Intelligence (AI) models, which can learn from human language, behavior and preferences. Although this practice enables significant advances in content personalization and network moderation, it also raises ethical questions, especially in relation to users' privacy and the appropriate use of their personal data. This article examines the methods of collecting and processing social media data, the main models and algorithms applied, and the ethical implications associated with the use of this information. It concludes with an analysis of the impact of this practice on users' privacy and recommendations for a more ethical and transparent approach.

Keywords: artificial intelligence; AI training; privacy; social networks; navigation data.

INTRODUÇÃO

Com a ascensão das redes sociais, plataformas como Instagram, Meta (Facebook) e X (anteriormente Twitter) passaram a coletar e armazenar grandes volumes de dados, que refletem

¹ Graduanda em Direito pela FMU – SP.

as interações humanas em ambiente digital. Estes dados incluem postagens, curtidas, compartilhamentos e comentários, e oferecem um cenário vasto e diversificado para o treinamento de modelos de IA. Essa prática tem o potencial de gerar avanços tecnológicos em áreas como personalização de conteúdo, detecção de discurso de ódio, análise de sentimentos e recomendação de produtos. No entanto, o uso de postagens para treinamento de IA suscita uma série de questões éticas, sendo a privacidade dos usuários uma das mais prementes.

Este artigo busca aprofundar a análise dos métodos e práticas que envolvem o uso de postagens de redes sociais para o treinamento de modelos de IA, abordando as tecnologias envolvidas, as aplicações e os impactos na privacidade dos indivíduos.

DISCUSSÕES E RESULTADOS

Metodologia de Treinamento de IA com Dados de Redes Sociais

O treinamento de IA com dados provenientes de redes sociais exige etapas estruturadas e cuidadosas de coleta, processamento, anonimização e balanceamento dos dados. A seguir, destacam-se essas etapas e suas características:

Coleta de Dados

A coleta de dados de redes sociais pode ser realizada por meio de interfaces de programação de aplicações (APIs) oferecidas por algumas redes, como Twitter e Instagram utilizando:

- **Filtro de tópicos e relevância:** O uso de palavras-chave, hashtags e outras características de filtragem para assegurar que o conteúdo coletado seja relevante para os objetivos do treinamento.
- **Frequência e volume de dados:** Redes sociais produzem grande volume de dados diariamente, exigindo métodos de coleta que possam manejar e armazenar eficientemente essa massa de informações.

Pré-Processamento de Dados

Os dados de redes sociais são predominantemente não estruturados e apresentam alto nível de informalidade, o que exige processamento para remover ruídos e ajustar o conteúdo. Esse pré-processamento envolve:

- **Limpeza e normalização:** Remoção de ruídos, como URLs, emojis e dados irrelevantes.

- Tokenização e vetorização: Estruturação dos textos para que possam ser processados por modelos de linguagem natural, convertendo as palavras em vetores de números que representam semanticamente os termos.

Anonimização e Privacidade

Devido à natureza pública e pessoal das redes sociais, o processo de anonimização dos dados é fundamental para garantir a privacidade dos usuários. Técnicas comuns incluem:

- Remoção de identificadores pessoais: Remoção de nomes, localizações e outros dados que possam identificar o usuário diretamente.
- Generalização de dados: Técnicas que substituem informações específicas por categorias mais amplas, reduzindo o risco de identificação sem perder contexto para o treinamento da IA.

Balanceamento e Diversidade de Dados

Para evitar vieses, é necessário que o conjunto de dados represente uma variedade demográfica e cultural. Em redes sociais, onde há uma pluralidade de grupos, o balanceamento de dados ajuda a minimizar o risco de enviesamento nos modelos de IA.

Aplicações de Modelos de IA Treinados com Dados de Redes Sociais

Redes sociais têm sido utilizadas para o desenvolvimento de modelos de IA em várias aplicações, algumas das quais incluem:

Análise de Sentimentos e Previsão de Tendências

Modelos de IA treinados em redes sociais são amplamente utilizados para interpretar o sentimento do público em relação a determinados tópicos. Isso é especialmente útil para empresas que monitoram a percepção dos consumidores sobre produtos, marcas ou eventos.

Moderação e Segurança de Conteúdo

Redes como Meta e X utilizam IA para identificar e moderar conteúdos nocivos, incluindo discurso de ódio, conteúdo violento e desinformação. Modelos de IA analisam postagens em tempo real, removendo conteúdos que infrinjam as políticas da plataforma e promovendo um ambiente mais seguro.

Algoritmos de Recomendação e Personalização

Esses modelos são essenciais para o fornecimento de conteúdo personalizado aos usuários. Utilizando o histórico de interações e preferências de cada indivíduo, a IA recomenda postagens, páginas e até produtos, aumentando a retenção e o engajamento na plataforma.

Publicidade Direcionada e Monetização

Os dados coletados também são utilizados para aprimorar a segmentação de anúncios, permitindo que redes sociais direcionem anúncios mais alinhados ao perfil de interesses de cada usuário. Essa prática potencializa a eficácia dos anúncios e, por consequência, os lucros da plataforma. No entanto, para que essa segmentação ocorra, é necessário um uso intensivo dos dados dos usuários, o que gera preocupações sobre privacidade e direitos autorais.

A publicidade direcionada, é uma estratégia amplamente utilizada por plataformas digitais, como o Instagram, Facebook e X, e se baseia na coleta e análise de dados dos usuários para personalizar anúncios levantando uma série de preocupações sobre a privacidade dos usuários. A seguir, exploramos os principais aspectos dessa questão:

Coleta Excessiva de Dados

Para oferecer publicidade direcionada, as plataformas precisam coletar uma vasta quantidade de dados dos usuários. Isso inclui não apenas informações explícitas, como dados demográficos e interesses informados, mas também dados implícitos, que são inferidos a partir do comportamento online, como:

- Interações com postagens: Curtidas, comentários e compartilhamentos ajudam a identificar interesses específicos.
- Navegação na web: O rastreamento de páginas visitadas e tempo gasto em diferentes conteúdos revela padrões de consumo e preferências.
- Atividades fora da plataforma: Muitos anunciantes utilizam cookies e pixel tracking para monitorar o comportamento do usuário em outros sites, o que pode ser combinado com dados de redes sociais.

Essa coleta massiva pode ser vista como uma invasão à privacidade, especialmente quando os usuários não têm ciência clara sobre quais dados estão sendo coletados e como serão usados.

Consentimento e Transparência

A publicidade direcionada geralmente depende do consentimento do usuário para a coleta de dados. Contudo, muitos usuários não leem ou compreendem as políticas de privacidade e os termos de serviço, o que pode resultar em um consentimento informado inadequado.

- **Falta de clareza:** Os usuários podem não entender quais informações estão sendo coletadas, como estão sendo usadas e por que motivos. Isso gera uma falta de transparência que prejudica a confiança na plataforma.

- **Consentimento implícito:** Em alguns casos, o consentimento é considerado implícito, levando à coleta de dados sem uma autorização explícita. Essa prática pode ser controversa, especialmente quando os usuários não estão cientes de que estão optando por compartilhar suas informações.

Reidentificação e Risco de Exposição

Mesmo quando os dados são anonimizados, existe o risco de reidentificação. Isso pode ocorrer quando dados coletados de várias fontes são combinados, permitindo que anunciantes ou terceiros identifiquem indivíduos específicos.

- **Rastreamento de atividades:** Se um usuário interage com múltiplas plataformas que compartilham informações entre si, suas atividades podem ser traçadas de volta a ele, comprometendo sua privacidade.

- **Vulnerabilidades de segurança:** Dados coletados podem ser vulneráveis a vazamentos ou acessos não autorizados, expondo informações pessoais que os usuários prefeririam manter em sigilo.

Manipulação e Controle

A personalização dos anúncios pode levar a uma forma de manipulação comportamental. Quando os usuários são constantemente expostos a anúncios que refletem suas preferências, isso pode moldar suas decisões de compra e influenciar seu comportamento sem que eles estejam cientes.

- **Criação de "câmaras de eco":** Anúncios personalizados podem reforçar crenças e preferências existentes, limitando a exposição a novas ideias e produtos, o que pode resultar em uma visão distorcida da realidade.

- **Impacto em decisões sensíveis:** Isso é particularmente preocupante em áreas como política e saúde, onde a manipulação da informação pode influenciar decisões importantes.

Regulamentação e Proteção de Dados

Diante das preocupações sobre privacidade, diversas jurisdições implementaram regulamentações, como o GDPR na União Europeia e a LGPD no Brasil, que visam proteger os dados dos usuários.

- **Direito à privacidade:** Essas leis oferecem aos usuários maior controle sobre suas informações, permitindo que eles solicitem a exclusão de dados, revisem as informações que uma plataforma possui sobre eles e optem por não participar de atividades de rastreamento.
- **Responsabilidade das empresas:** As plataformas são obrigadas a garantir que as práticas de coleta e uso de dados sejam transparentes e respeitem as preferências dos usuários. O não cumprimento pode resultar em penalidades severas.

Desafios Éticos e Impacto na Privacidade

Embora as redes sociais ofereçam dados valiosos para o treinamento de IA, a utilização dessas informações levanta desafios éticos, especialmente no que se refere à privacidade e à conformidade com regulamentações de proteção de dados, como o GDPR e a LGPD.

Questões de Privacidade e Consentimento

A coleta de dados de redes sociais para IA frequentemente ocorre de forma indireta, sem que os usuários tenham ciência ou consentam explicitamente para o uso de seus dados para treinamento. Essa prática viola o direito à privacidade, compromete a transparência da plataforma com seus usuários.

Viés e Discriminação

Modelos de IA treinados com dados de redes sociais podem refletir vieses demográficos e culturais inerentes às interações nas plataformas, uma vez que certos grupos estão mais representados que outros. Isso pode gerar modelos com respostas enviesadas, especialmente em questões sociais e políticas, reforçando estereótipos e discriminações.

Uso de Dados para Vigilância e Controle

Os dados das redes sociais treinam modelos que podem ser utilizados para vigilância social e controle de comportamento, com riscos de invasão de privacidade, monitoramento sem consentimento e potencial manipulação de decisões de usuários.

Perspectivas para Mitigação de Riscos e Futuras Direções

Dado o potencial impacto negativo do uso de redes sociais para treinamento de IA sobre a privacidade, diversas abordagens podem ser adotadas para reduzir os riscos e assegurar um uso mais ético dos dados:

Adoção de Normas Transparentes e Consentimento Esclarecido

É necessário que as redes sociais adotem políticas de privacidade que esclareçam como os dados são utilizados e obtenham o consentimento dos usuários de forma explícita para o uso de informações pessoais no treinamento de IA.

Controle e Auditoria de Viés nos Modelos de IA

Modelos treinados com dados de redes sociais devem passar por auditorias para detecção e correção de vieses. Métodos de equalização de dados e controle de variáveis demográficas podem ser empregados para garantir maior imparcialidade e representatividade.

Regulação e Compliance com Regulamentações de Privacidade

A implementação de regulamentações como o GDPR e a LGPD torna-se fundamental, exigindo que as plataformas se adequem a um padrão de proteção de dados, respeitando os direitos dos usuários. As redes sociais devem cumprir rigorosamente as normas de proteção de dados, sujeitando-se a auditorias e sanções em caso de falhas.

CONCLUSÃO

O treinamento de IA a partir de dados de redes sociais oferece benefícios expressivos em termos de personalização de conteúdo, moderação e aprimoramento da experiência digital. No entanto, o uso de informações de usuários requer uma abordagem ética que respeite a privacidade e os direitos individuais. Questões como anonimização, vieses e compliance com regulamentações de proteção de dados são centrais para o desenvolvimento responsável desses modelos. À medida que a IA evolui, é imperativo que as plataformas adotem práticas transparentes e responsáveis, proporcionando uma experiência digital segura e respeitosa com os dados de seus usuários.

REFERÊNCIAS

CÉSAR, P. **Conheça as Etapas do Pré-Processamento de dados - DataGeeks**. Disponível em: <https://www.datageeks.com.br/pre-processamento-de-dados/>. Acesso em: 27 out. 2024.

DA CNN. IA generativa: especialista explica à CNN como é o treinamento da ferramenta. Disponível em: <https://www.cnnbrasil.com.br/tecnologia/ia-generativa-especialista-explica-a-cnn-como-e-o-treinamento-da-ferramenta/>. Acesso em: 27 out. 2024.

ETAIN. LGPD Cookies: o guia completo | SUA IMPRENSA. Disponível em: <https://suaimpresa.com.br/blog/lgpd-cookies-o-guia-completo/>. Acesso em: 27 out. 2024.

FOTOS. Instagram e Facebook usam suas fotos e textos para treinar IA: por que a prática está sendo questionada? Disponível em: <https://g1.globo.com/tecnologia/noticia/2024/07/01/instagram-e-facebook-usam-suas-fotos-e-textos-para-treinar-ia-por-que-a-pratica-esta-sendo-questionada.ghtml>. Acesso em: 27 out. 2024.

HTTPS://GBASTOS811.JUSBRASIL.COM.BR. Proteção de dados nas redes sociais | Jusbrasil. Disponível em: <https://www.jusbrasil.com.br/artigos/protecao-de-dados-nas-redes-sociais/1810106087>. Acesso em: 27 out. 2024.

LGPD: como lidar com a proteção de dados nas mídias sociais? | Becker Direito Empresarial. Disponível em: <https://www.direitoempresarial.com.br/lgpd-como-lidar-com-a-protecao-de-dados-nas-midias-sociais>. Acesso em: 27 out. 2024.

META. Meta não pode usar dados de usuários de redes sociais para treinar IA - Migalhas. Disponível em: <https://www.migalhas.com.br/quentes/410502/meta-nao-pode-usar-dados-de-usuarios-de-redes-sociais-para-treinar-ia>. Acesso em: 27 out. 2024.

O que é: Pré-Processamento De Dados. Disponível em: <https://estatisticafacil.org/glossario/o-que-e-pre-processamento-de-dados-guia-completo/>. Acesso em: 27 out. 2024.

VEJA. Governo manda Meta interromper treinamento de IA com dados de usuários. Disponível em: <https://veja.abril.com.br/coluna/maquiavel/governo-manda-meta-interromper-treinamento-de-ia-com-dados-de-usuarios>. Acesso em: 27 out. 2024.