abpi.empauta.com

Associação Brasileira da Propriedade Intelectual
Clipping da imprensa

Brasília, 08 de julho de 2024 às 07h53
Seleção de Notícias

abpi.empauta.com

Inteligência Artificial apreende 32 bilhões de produtos piratas de clubes

Assine UOL **Pirataria** dá prejuízo para clubes de futebol Imagem: Talyta Vespa/UOLUm sistema de Inteligência Artificial está atuando contra a **pirataria** que atinge clubes de futebol em vendas online. Segundo levantamento feito pela Bianchini Advogados, 32 bilhões de produtos ilegais já foram apreendidos com a atuação dessa tecnologia em diferentes sites de e-commerce.

O sistema monitora 24 horas por dia e sete dias por semana potenciais vendedores piratas, identifica produtos que não poderiam usar as marcas de times de futebol e entrega para um time de especialistas que aí entra com ações na Polícia e também na Receita Federal.

"Depois da pandemia, a **pirataria** sofreu uma mudança substancial para o ambiente virtual. A exposição que antes era só na rua agora está na internet. Então passamos a utilizar a inteligência artificial e começamos a rastrear esse tipo de coisa. E não é só camisa de futebol. Caneca, caneta, tábua de churrasco, lápis, maquete de estádio, baralho, taco de sinuca Olha, são tantos produtos que ostentam a marca do clube e não geram os devidos royalties aos detentores da marca", explica Ricardo Bianchini, sócio do escritório.

"A gente identifica e, no plano físico, conseguimos ir ao local dos **produtos** piratas em parceria com a polícia e apreende a mercadoria para, posteriormente, fazer a destruição. Em sites, a gente consegue acionar o vendedor, quando conseguimos identificar de primeira, ou até mesmo notificamos a plataforma que estiver fazendo a venda. Conseguimos derrubar, inclusive, páginas internacionais", completou.

Hoje, Santos, Bahia e Corinthians trabalham com essa tecnologia, mas outros clubes como o São Paulo, por exemplo, também já tiveram esse tipo de ajuda.

Os clubes ainda sofrem com quem já teve a autorização para licenciar produtos vencidas, com os que burlam a quantidade limitada mesmo sendo licenciado e até mesmo com sites que praticam golpes fingindo serem os sites oficiais dos clubes, mas que não revendem produtos em nenhum momento.

"A gente começa atribuindo de R$ 20 mil a R$ 25 mil reais e o critério é a gravidade da violação da marca. A gente tem uma compreensão social que a marca do clube é de domínio público, que tudo bem falsificar uma camisa pirata, que a marca é de todo mundo. Mas não é assim. Já cheguei a pegar uma loja que era especializada em vender produtos de festas com os escudos dos times, sem nenhum tipo de autorização dos times", finalizou.

Hackers miram em empresas de inteligência artificial para furto de dados

Fonte: CenárioMT

Não precisa se preocupar se suas conversas secretas com o ChatGPT vazaram no recente ataque aos sistemas da OpenAI. O hack em si, embora preocupante, parece ter sido superficial, mas serve como um lembrete de que as empresas de IA se tornaram, em pouco tempo, alvos tentadores para hackers.

O jornal The New York Times relatou o ataque com mais detalhes, após o ex-funcionário da OpenAI, Leopold Aschenbrenner, ter dado a entender em um podcast recentemente. Ele o chamou de "grande incidente de segurança", mas fontes não identificadas da empresa disseram ao Times que o hacker só conseguiu acessar um fórum de discussão de funcionários. (O repórter tentou contatar a OpenAI para confirmação e comentário.)

Nenhuma violação de segurança deve ser tratada como trivial, e espionar conversas internas sobre desenvolvimento de IA certamente tem valor. Mas está longe de ser um hacker obtendo acesso a sistemas internos, modelos em progresso, roteiros secretos e assim por diante.

No entanto, isso deveria nos assustar de qualquer maneira, e não necessariamente por causa da ameaça da China ou de outros países nos superando na corrida armamentista da IA. O simples fato é que essas empresas de IA se tornaram guardiãs de uma enorme quantidade de dados valiosos.

O que os hackers buscam

Vamos falar sobre três tipos de dados que a OpenAI e, em menor medida, outras empresas de IA criaram ou têm acesso: dados de treinamento de alta qualidade, interações em massa com usuários e dados de clientes.

É incerto quais dados de treinamento eles possuem exatamente, porque as empresas são extremamente sigilosas sobre seus tesouros. Mas é um erro pensar que eles são apenas grandes pilhas de dados da web coletados automaticamente. Sim, eles usam coletores da web ou conjuntos de dados como o Pile, mas é uma tarefa gigantesca moldar esses dados brutos em algo que possa ser usado para treinar um modelo como o GPT-4. Um grande número de horas de trabalho humano é necessário para fazer isso - só pode ser parcialmente automatizado.

Alguns engenheiros de aprendizado de máquina especulam que, de todos os fatores que influenciam a criação de um modelo de linguagem grande (ou talvez qualquer sistema baseado em transformadores), o mais importante é a qualidade do conjunto de dados. É por isso que um modelo treinado no Twitter e no Reddit nunca será tão eloquente quanto um treinado em todas as obras publicadas do século passado. (E provavelmente por que a OpenAI supostamente usou fontes legalmente questionáveis, como livros protegidos por **direitos** autorais, em seus dados de treinamento, uma prática que eles afirmam ter abandonado.)

Portanto, os conjuntos de dados de treinamento que a OpenAI construiu são de enorme valor para os concorrentes, desde outras empresas a estados adversários e reguladores aqui nos Estados Unidos. As autoridades não gostariam de saber exatamente quais dados estão sendo usados e se a OpenAI tem sido honesta sobre isso?

Mas talvez ainda mais valiosa seja a enorme quantidade de dados de usuários da OpenAI - provavelmente bilhões de conversas com o ChatGPT sobre centenas de milhares de tópicos. Assim como os dados de pesquisa já foram a chave para entender a mente coletiva da web, o ChatGPT está na cola da pulsação de uma população que pode não ser tão ampla quanto o universo de usuários do Google, mas fornece muito mais profundidade. (Caso você não saiba, a menos que desative a opção, suas conversas estão sendo usadas para dados de treinamento.)

Desde compras até dados pessoais

No caso do Google, um aumento nas pesquisas por "ar-condicionados" indica que o mercado está esquentando um pouco. Mas esses usuários não conversam sobre o que querem, quanto estão dispostos a gastar, como é a casa deles, fabricantes que querem evitar e assim por diante. Você sabe que isso é valioso porque o próprio Google está tentando converter seus usuários para fornecer essas mesmas informações, substituindo as pesquisas por interações com IA!

Imagine quantas conversas as pessoas tiveram com o ChatGPT e quão útil essa informação é, não apenas para desenvolvedores de IAs, mas para equipes de marketing, consultores, analistas é uma mina de ouro.

A última categoria de dados talvez seja a de maior valor no mercado aberto: como os clientes estão realmente usando a IA e os dados que eles próprios forneceram aos modelos.

Centenas de grandes empresas e incontáveis empresas menores usam ferramentas como as APIs da OpenAI e da Anthropic para uma variedade igualmente grande de tarefas. E para que um modelo de linguagem seja útil para elas, geralmente ele precisa ser ajustado ou ter acesso a seus próprios bancos de dados internos.

Isso pode ser algo prosaico como planilhas de orçamento antigas, registros de pessoas (para torná-las mais fáceis de pesquisar, por exemplo) ou tão valioso quanto.

Índice remissivo de assuntos