

Análise do caso: Viés no recrutamento (Amazon)

Contexto do caso:

“A ferramenta experimental de contratação da Amazon utilizou inteligência artificial para atribuir aos candidatos a vagas de emprego na companhia pontuações que variavam de uma a cinco estrelas. Mas em 2015 a empresa percebeu que seu novo sistema não classificava candidatos para empregos de desenvolvedores de software e outros cargos técnicos de maneira neutra em termos de gênero.

Isso porque os modelos de computador da Amazon foram treinados para examinar os candidatos observando padrões em currículos enviados à empresa durante um período de 10 anos. A maioria veio de homens, um reflexo do domínio masculino em toda a indústria de tecnologia.

Como consequência, o sistema aprendeu a favorecer candidatos do sexo masculino e penalizar currículos que incluíam a palavra “women’s”, além de baixar as pontuações de formados em instituições exclusivamente femininas. A empresa de Seattle encerrou o projeto no início de 2018 porque os executivos perderam a confiança na ferramenta e não vislumbravam meios eficazes de corrigir o problema.”

Leia mais em: <https://forbes.com.br/last/2018/10/amazon-desiste-de-ferramenta-secreta-de-recrutamento/>

Analisando o caso, é possível compreender melhor algumas questões sobre o viés. O conjunto de dados utilizado para treinar a IA foi constituído, majoritariamente, de currículos de candidatos homens, refletindo os padrões de dominância dentro da área de TI na época. O algoritmo, portanto, aprendeu a replicar e reforçar desigualdades de gênero, penalizando termos como “women’s”. Com isso, a tecnologia reduziu as oportunidades de emprego para mulheres dentro da companhia, aprofundando desigualdades já existentes.

O sistema funcionava como uma black box, pois não havia explicabilidade clara sobre as decisões da IA. Isso dificultava a realização de auditorias internas e comprometia a transparência do processo, resultando na perda de confiança dos

gestores e inviabilizando a correção do problema.

No mercado de trabalho, o sistema intensificou as barreiras para que mulheres ingressassem em cargos de tecnologia, sem que elas pudessem compreender ou contestar o motivo da exclusão de seus currículos. Esse cenário afetou diretamente princípios de não discriminação e autonomia informacional, previstos inclusive na LGPD.

A equipe de desenvolvimento poderia ter conduzido auditorias éticas e testes de viés antes da implementação do projeto, o que teria possibilitado a identificação e correção do problema. Os princípios de “Ethical AI by Design” que poderiam ter sido aplicados incluem fairness (justiça), accountability (responsabilidade) e transparency (transparência). Além disso, a conformidade com a LGPD seria essencial para a proteção dos dados dos candidatos e para assegurar o direito à explicação sobre decisões automatizadas.

Meu posicionamento sobre o caso reforça a decisão da própria empresa. O sistema deveria ser descontinuado. No entanto, considerando as tecnologias aplicáveis à área de R&S, ele poderia ser redesenhado e aprimorado, desde que integrado a princípios de ética e governança.

Minhas recomendações são: implementação de auditorias contínuas nos dados do algoritmo, aplicação de mecanismos de explicabilidade nas decisões tomadas e criação de comitês de governança ética multidisciplinares, envolvendo TI, RH e Jurídico, para supervisionar o desenvolvimento e corrigir discrepâncias e desigualdades sempre que identificadas.