

Relatório da experiência prática 3 – Análise de caso onde a IA causou um problema ético

Autor: Vicente Philipe Carvalho Magnani

Introdução

O relatório tem como objetivo analisar o caso em que um software COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), usado por juízes em vários estados americanos para prever a probabilidade de um réu cometer crimes futuros era significativamente mais propenso a rotular réus negros como de "alto risco" de reincidência. E após isso irei dar minha opinião e recomendações.

Para que ele serve?

O COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) é um algoritmo usado por juízes em vários estados americanos para prever a probabilidade de um réu cometer crimes futuros, a ideia era ajudar nas decisões de sentenças, fianças e liberdade condicional.

O problema étnico do COMPAS

Durante uma investigação de 2016 feita pela ProPublica revelou um viés alarmante no sistema, a análise mostrou que o algoritmo era significativamente mais propenso a rotular réus negros como de "alto risco" de reincidência, mesmo quando eles não reincidiam. Em contrapartida, réus brancos eram classificados como de "baixo risco", mesmo que tivessem antecedentes criminais piores.

Viés contra réus negros: Os réus negros tinham o dobro da probabilidade de serem classificados erroneamente como de alto risco comparado aos réus brancos.

Viés a favor de réus brancos: Os réus brancos tinham maior probabilidade de serem classificados como de baixo risco em comparação aos réus negros.

Por que o viés aconteceu?

O problema ético não estava na intenção do sistema, mas nos dados que ele usava para aprender. O COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) foi treinado com dados históricos de prisões, que refletiam décadas de desigualdade social e racismo estrutural no sistema de justiça criminal dos EUA. Por exemplo, a polícia historicamente patrulha bairros de minorias com mais frequência, o que leva a um número desproporcional de prisões nessas áreas. O algoritmo, sem entender o contexto social, interpretou esse padrão como um sinal de que morar em certos bairros ou ter certas características significava um maior risco de reincidência.

As consequências éticas

Este caso levantou questões éticas cruciais sobre o uso de IA em decisões que afetam a vida das pessoas. O COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) mostrou que um sistema de IA, mesmo que seja neutro em sua programação, pode perpetuar e até amplificar os preconceitos existentes na sociedade se os dados de treinamento não forem cuidadosamente examinados. A falta de transparência sobre como o algoritmo chegou a suas conclusões também foi um grande problema, tornando impossível para os réus contestarem a decisão.

Por que o COMPAS continua em uso?

Validação de suas previsões: Os defensores do COMPAS argumentam que o sistema, apesar de suas falhas, é mais preciso na previsão de reincidência do que a avaliação de um juiz ou oficial de liberdade condicional baseada apenas na intuição humana.

Economia de recursos: Em muitos casos, esses algoritmos ajudam a acelerar o processo judicial, o que pode ser visto como uma forma de otimizar o tempo e os recursos do sistema de justiça.

Falta de alternativas: Não há um consenso sobre qual seria a melhor alternativa para prever o risco de reincidência, e muitos sistemas judiciais continuam a recorrer a ferramentas como o COMPAS para auxiliar suas decisões.

O que mudou de 2016 para 2025?

Maior escrutínio: Houve um aumento do debate público e acadêmico sobre a ética e a equidade no uso de IA no sistema de justiça.

Regulamentação e transparência: Algumas leis foram propostas para exigir mais transparência em como esses algoritmos funcionam e para garantir que eles sejam auditados regularmente em busca de vieses.

Desenvolvimento de novos modelos: Pesquisadores e desenvolvedores estão trabalhando em novos modelos de avaliação de risco que buscam mitigar o viés e oferecer maior transparência.

Conclusão

embora o COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) ainda seja utilizado, ele opera em um contexto muito mais rigoroso e sob um escrutínio público e regulatório muito maior do que antes da controvérsia vir à tona.

1º Opinião e sugestão: O Sistema é útil, mas precisa de melhorias e supervisão humana

Acredito que a proibição do sistema COMPAS não seja a medida mais adequada. Em vez disso, defendo a necessidade de aprimoramentos contínuos no algoritmo, considerando que o sistema foi treinado com dados históricos de prisões, que refletem as desigualdades e os vieses existentes no próprio sistema de justiça criminal, o problema subjacente reside na estrutura social e não apenas na tecnologia. Portanto, em vez de descartar a ferramenta, o foco deve ser em corrigir os dados de entrada e refinar o algoritmo para que ele se torne mais justo e equitativo.

2º opinião e sugestão: melhorar a sociedade como um todo

O ideal seria um mundo onde a tecnologia de IA, como a utilizada no COMPAS, pudesse ser treinada com dados que refletissem uma sociedade mais justa e equitativa. Se os dados históricos de prisões já fossem livres de preconceitos e racismo, o algoritmo não teria como perpetuar a discriminação. Isso significa que o caminho para o aprimoramento da IA passa, necessariamente, pelo esforço coletivo de melhorar a sociedade como um todo, tornando-a mais adequada e aceitável para

todos os seres humanos, sem preconceitos e sem racismos, a tecnologia nesse cenário, deixaria de ser um espelho de nossas falhas e se tornaria uma ferramenta poderosa para construir um futuro mais justo. O COMPAS nos mostra que, antes de programar o código, precisamos programar as nossas próprias atitudes.