



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I - CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE COMPUTAÇÃO
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

MARCOS LUCAS SOUSA ALVES

**A INFLUÊNCIA DOS VIESES ALGORÍTMICOS NAS DECISÕES DE
JULGAMENTO: o caso do COMPAS**

**CAMPINA GRANDE – PB
2023**

MARCOS LUCAS SOUSA ALVES

**A INFLUÊNCIA DOS VIESES ALGORÍMICOS NAS DECISÕES DE
JULGAMENTO: o caso do COMPAS**

Trabalho de Conclusão de Curso (Artigo) apresentado ao Departamento do Curso Ciência da Computação da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Bacharel em Ciências da Computação.

Orientador: Prof. Dr. Wellington Candeia de Araújo.

**CAMPINA GRANDE - PB
2023**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

A474i Alves, Marcos Lucas Sousa.

A influência dos vieses algorítmicos nas decisões de julgamento [manuscrito] : o caso do COMPAS / Marcos Lucas Sousa Alves. - 2023.

13 p.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Computação) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2023.

"Orientação : Prof. Dr. Wellington Candeia de Araujo, Coordenação do Curso de Computação - CCT. "

1. Viés Algorítmico. 2. Decisão de julgamento.
3. ProPublica. I. Título

21. ed. CDD 005.3

MARCOS LUCAS SOUSA ALVES

**A INFLUÊNCIA DOS VIESES ALGORÍTMICOS NAS DECISÕES DE
JULGAMENTO: o caso do COMPAS**

Trabalho de Conclusão de Curso (Artigo)
apresentado ao Departamento do Curso
Ciência da Computação da Universidade
Estadual da Paraíba, como requisito
parcial à obtenção do título de Bacharel em
Ciências da Computação.

Aprovado em: 29/06/2023

BANCA EXAMINADORA



Prof. Dr. **Wellington Candeia de Araújo** (Orientador)
Universidade Estadual da Paraíba (UEPB)

Documento assinado digitalmente



FRANCISCO ANDERSON MARIANO DA SILVA

Data: 10/07/2023 12:24:34-0300

Verifique em <https://validar.iti.gov.br>

Prof. Me. **Francisco Anderson Mariano Da Silva** (CCEA/UEPB)
Universidade Estadual da Paraíba (UEPB)



Prof. Me. **Vinicius Reuteman Feitosa Alves de Andrade** (CCEA/UEPB)
Universidade Estadual da Paraíba (UEPB)

A minha mãe e minha esposa, pela
dedicação, companheirismo e amizade,
DEDICO.

SUMÁRIO

1	INTRODUÇÃO	6
2	REFERENCIAL TEÓRICO	6
2.1	Conceitos de algoritmo e aprendizado de máquina	7
2.2	Viés Algorítmico	7
2.3	COMPAS e seu funcionamento	8
2.4	Principais vieses algorítmicos que influenciam as decisões de julgamento	9
2.5	Possíveis soluções para reduzir ou eliminar os vieses algorítmicos que afetam as decisões de julgamento.....	9
3	METODOLOGIA	10
4	RESULTADOS E DISCUSSÕES	11
5	CONSIDERAÇÕES FINAIS	12
	REFERÊNCIAS.....	12

A INFLUÊNCIA DOS VIESES ALGORÍTMICOS NAS DECISÕES DE JULGAMENTO: o caso do COMPAS

THE INFLUENCE OF ALGORITHMIC BIASES ON JUDICIAL DECISIONS: the case of COMPAS

Marcos Lucas Sousa Alves¹

RESUMO

O viés algorítmico é uma distorção sistemática que pode ocorrer quando algoritmos são treinados em dados que refletem preconceitos e desigualdades presentes na sociedade. No caso do COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), vieses têm sido identificados, resultando em disparidades nas decisões de julgamento. Nesse contexto, torna-se fundamental analisar a influência dos vieses algorítmicos nas decisões de julgamento, para atenuar ou mitigar os mesmos. Diferentes soluções têm sido propostas, como melhorar os conjuntos de dados utilizados, aumentar a transparência e auditoria dos algoritmos e implementar medidas corretivas. No entanto, a discussão sobre vieses algorítmicos no COMPAS é complexa e controversa, com alguns argumentando que os algoritmos podem levar a decisões mais imparciais, enquanto outros questionam sua objetividade. A análise de conjuntos de dados compilados pela ProPublica revelou evidências de vieses raciais, de gênero e socioeconômicos no COMPAS, reforçando a necessidade de aprimoramentos no sistema. Abordar o viés algorítmico no COMPAS é crucial para ajudar a garantir decisões de julgamento mais justas e equitativas no sistema de justiça criminal.

Palavras-chave: viés algorítmico; decisão de julgamento; ProPublica.

ABSTRACT

The algorithmic bias is a systematic distortion that can occur when algorithms are trained on data that reflects prejudices and inequalities present in society. In the case of COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), biases have been identified, resulting in disparities in judgment decisions. In this context, it becomes crucial to analyze the influence of algorithmic biases on judgment decisions in order to mitigate or alleviate them. Different solutions have been proposed, such as improving the datasets used, increasing transparency and auditing of algorithms, and implementing corrective measures. However, the discussion on algorithmic biases in COMPAS is complex and controversial, with some arguing that algorithms can lead to more impartial decisions, while others question their objectivity. Analysis of datasets compiled by ProPublica has revealed evidence of racial, gender, and socioeconomic biases in COMPAS, reinforcing the need for system improvements. Addressing algorithmic bias in COMPAS is crucial to help ensure fair and equitable judgment decisions in the criminal justice system.

¹ Graduando em Ciência da Computação - Email: marcoslucas08@gmail.com; LinkedIn: <https://www.linkedin.com/in/marcoslucas/>

Keywords: algorithmic bias; judicial decision; ProPublica

1 INTRODUÇÃO

Com o aumento da adoção de sistemas algorítmicos nos últimos anos para prever o comportamento das pessoas, a fim de auxiliar as decisões de julgamento, coloca em evidência questões relacionadas à discriminação por motivo de raça, classe social e gênero. Embora o COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) tenha sido desenvolvido com o objetivo de auxiliar juízes na tomada de decisões mais justas e imparciais, estudos recentes revelam a presença de vieses algorítmicos que podem influenciar significativamente o resultado das sentenças.

Um estudo do ProPublica² sobre o sistema COMPAS, apontou evidências de que, quando utilizado para prever o risco de reincidência criminal, o algoritmo de previsão era significativamente mais propenso a classificar afro-americanos como mais propensos a reincidir do que os brancos (ANGWIN *et al.*, 2016). Por isso, o aprofundamento sobre os vieses algorítmicos em sistemas como o COMPAS, que se tornaram cada vez mais presentes em áreas como a segurança pública e a Justiça, é essencial para garantir a imparcialidade e a justiça das decisões judiciais.

Diante dessas preocupações, surge a necessidade de investigar a influência dos vieses algorítmicos nas decisões judiciais, especificamente no caso do COMPAS. Como objetivos específicos conceituar viés algorítmico, descrever o COMPAS e qual o seu funcionamento e identificar os principais vieses algorítmicos que influenciam as decisões de julgamento, discutindo os impactos desses vieses nas decisões tomadas pelos juízes. Ao compreender melhor esses vieses, será possível propor reflexões e medidas que visem atenuar e mitigar a discriminação algorítmica, promovendo assim um sistema de justiça mais equitativo e justo.

É importante ressaltar que a análise dos vieses no COMPAS não visa questionar a utilidade dos algoritmos na justiça criminal, mas sim garantir que essas ferramentas sejam desenvolvidas e aplicadas de forma ética, respeitando os direitos fundamentais e promovendo a igualdade perante a lei.

O presente trabalho se apresenta como segue: na seção dois tem-se o referencial teórico; na seção três a metodologia; na seção quatro apresentam-se os resultados do trabalho; por fim, na seção cinco as considerações finais e referências.

2 REFERENCIAL TEÓRICO

Nesta seção serão apresentados conceitos de diferentes autores sobre qual a influência dos vieses algorítmicos nas decisões de julgamento presentes no COMPAS. Mas antes que se inicie a apresentação dos referenciais teóricos associados ao tema e conceitos deste trabalho, é importante destacar os conceitos de algoritmo e aprendizado de máquina, para discutir sobre a importância na análise crítica do algoritmo utilizado no sistema de justiça criminal dos Estados Unidos.

² <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

2.1 Conceitos de algoritmo e aprendizado de máquina

O algoritmo pode ser entendido como uma sequência de instruções bem definidas, que permite a resolução de um determinado problema ou a realização de uma tarefa específica. Cormen (2013, p. 1) define algoritmo como “uma série de etapas para completar uma tarefa que é descrita de maneira precisa o bastante para que um computador possa realizá-la”. Os algoritmos são amplamente utilizados na ciência da computação e desempenham um papel fundamental no processamento de dados e na tomada de decisões.

O aprendizado de máquina, por sua vez, é uma área da inteligência artificial que se concentra no desenvolvimento de algoritmos capazes de aprender e tomar decisões com base em dados. Mitchell (1997) definiu o aprendizado de máquina como um processo em que um programa de computador é capaz de melhorar seu desempenho em uma determinada tarefa T , medida pela métrica de desempenho M , com base na experiência E adquirida. O aprendizado de máquina permite que os algoritmos sejam treinados com dados históricos para realizar tarefas específicas, como classificação, regressão e previsão.

No contexto do COMPAS e dos sistemas de pontuação de risco utilizados no sistema de justiça criminal, os algoritmos de aprendizado de máquina são empregados para analisar dados históricos e gerar pontuações de risco, com o objetivo de auxiliar as decisões de julgamento (ANGWIN *et al.*, 2016). No entanto, é importante ressaltar que esses algoritmos podem estar sujeitos a vieses e injustiças, o que pode resultar em tratamentos discriminatórios ou desiguais.

Compreender os conceitos de algoritmo e aprendizado de máquina é essencial para uma análise aprofundada dos vieses algorítmicos no COMPAS e em sistemas similares. Isso nos permite questionar como esses algoritmos são projetados, treinados e aplicados, considerando as possíveis influências de fatores como raça, gênero e outros aspectos socioeconômicos. A compreensão desses conceitos nos capacita a avaliar criticamente as soluções propostas e buscar abordagens que mitiguem os vieses algorítmicos, visando um sistema de justiça mais equitativo e imparcial.

2.2 Viés Algorítmico

O viés algorítmico é um fenômeno que ocorre quando um algoritmo produz resultados discriminatórios ou injustos, influenciados por características pessoais, como raça, gênero, etnia ou outras variáveis sensíveis (HOWARD e BORENSTEIN, 2018). É importante compreender que os algoritmos são desenvolvidos por seres humanos e, portanto, podem refletir os preconceitos e as desigualdades existentes na sociedade. Essa problemática tem recebido crescente atenção de pesquisadores e acadêmicos, dada a utilização cada vez mais frequente de algoritmos em áreas críticas, como o sistema de justiça criminal.

Segundo Angwin *et al.* (2016) afirmam que o viés algorítmico pode ocorrer devido a diferentes razões, incluindo o viés nos dados de treinamento utilizados para desenvolver o algoritmo, a formulação inadequada das variáveis de entrada ou a falta de representatividade dos grupos minoritários no processo de desenvolvimento do algoritmo. Esses fatores podem levar a decisões discriminatórias e injustas, perpetuando desigualdades no sistema de justiça.

Barocas e Selbst (2016), definem o viés algorítmico como "o viés em um algoritmo que resulta em tratamento desigual de indivíduos ou grupos, sem justificativa legítima". Eles ressaltam que esses vieses podem ser intencionais ou não intencionais, e podem ocorrer tanto por ação quanto por omissão. Os autores também argumentam que os vieses algorítmicos podem ser particularmente problemáticos no contexto da justiça criminal, onde decisões errôneas podem ter consequências significativas para os indivíduos e para a sociedade como um todo.

Portanto, o conceito de viés algorítmico refere-se à tendência sistemática de algoritmos produzirem resultados discriminatórios ou injustos, influenciados por características pessoais ou variáveis sensíveis. É importante reconhecer a presença desse fenômeno e buscar soluções que visem mitigar os vieses algorítmicos, a fim de garantir um sistema de justiça mais equitativo e imparcial.

2.3 COMPAS e seu funcionamento

O COMPAS é um sistema de pontuação de risco amplamente utilizado no sistema de justiça criminal dos Estados Unidos. Desenvolvido pela empresa Northpointe, o COMPAS utiliza algoritmos de aprendizado de máquina para analisar dados históricos e fornecer uma pontuação de risco que visa auxiliar juízes na tomada de decisões sobre sentenciamento e liberdade condicional.

Destacam-se aqui autores que têm contribuído para a descrição e compreensão do funcionamento do COMPAS. Dieterich *et al.* (2016) explicam que o COMPAS utiliza um conjunto diversificado de variáveis para gerar uma pontuação de risco. Essas variáveis podem incluir informações sobre o histórico criminal, idade, emprego, histórico de abuso de substâncias, entre outros. O algoritmo analisa essas variáveis e atribui uma pontuação que indica o nível de risco de reincidência do indivíduo.

Segundo Dressel e Farid (2018), o COMPAS utiliza um modelo estatístico baseado em regressão para estimar a probabilidade de reincidência. Esse modelo é treinado com base em dados históricos de indivíduos que passaram pelo sistema de justiça criminal, sendo esses dados utilizados para identificar padrões e construir um modelo preditivo.

A organização jornalística ProPublica compilou um conjunto de dados abrangente, contendo informações sobre cerca de 10 mil indivíduos em Broward County, Flórida, que foram avaliados pelo COMPAS (ANGWIN *et al.*, 2016). Esse conjunto de dados tornou-se uma referência na análise dos vieses algorítmicos presentes no sistema.

Ao analisar o conjunto de dados da ProPublica, foram identificadas disparidades significativas na precisão das previsões do COMPAS entre diferentes grupos raciais. Em um artigo publicado pela ProPublica (ANGWIN *et al.*, 2016), foi constatado que o COMPAS tendia a classificar erroneamente pessoas negras com uma maior taxa de falsos positivos, ou seja, atribuindo a elas um risco mais alto do que o real, em comparação com indivíduos brancos. Essa disparidade levanta preocupações sobre a equidade do sistema e a possibilidade de discriminação algorítmica.

Além disso, Dressel e Farid (2018) realizaram uma análise detalhada do conjunto de dados da ProPublica, questionando a validade das variáveis utilizadas pelo COMPAS para fazer suas previsões. Eles argumentam que certas variáveis, como informações sobre condenações anteriores, podem servir como proxies de

fatores socioeconômicos e raciais, ampliando assim os vieses presentes no sistema.

2.4 Principais vieses algorítmicos que influenciam as decisões de julgamento

Estudos têm revelado a presença de vieses algorítmicos no COMPAS que podem influenciar significativamente as decisões de julgamento. Autores têm contribuído para a identificação e compreensão desses vieses, destacando suas consequências no sistema de justiça criminal.

Um dos principais vieses algorítmicos identificados no COMPAS é o viés racial. Angwin *et al.* (2016) demonstram em seu estudo que as pontuações de risco geradas pelo COMPAS são influenciadas por características demográficas, como raça, o que levanta sérias preocupações sobre a discriminação algorítmica. Segundo eles, indivíduos pertencentes a grupos minoritários podem receber pontuações de risco mais altas, mesmo quando têm perfis criminais semelhantes aos de indivíduos brancos. Essa disparidade resultante do viés racial pode levar a um tratamento desigual e injusto dentro do sistema de justiça.

Além do viés racial, outro viés algorítmico identificado é o viés de gênero. Dieterich *et al.* (2016) destacam que o COMPAS apresenta uma tendência de superestimar o risco de reincidência para indivíduos pertencentes a grupos minoritários, enquanto subestima o risco para indivíduos brancos, independentemente do gênero. Essa disparidade de tratamento com base no gênero pode resultar em decisões desiguais e injustas, afetando a vida de milhares de pessoas.

A falta de consideração de fatores contextuais relevantes também é apontada como um viés algorítmico no COMPAS. Dressel e Farid (2018) argumentam que o sistema pode não levar em conta circunstâncias individuais, como ambiente social, histórico pessoal e condições socioeconômicas. Esses fatores podem ter um impacto significativo na probabilidade de reincidência de um indivíduo, e a omissão dessas informações pode levar a avaliações imprecisas e injustas do risco de reincidência.

Esses vieses algorítmicos têm implicações importantes para a equidade e a imparcialidade do sistema de justiça. A presença desses vieses pode reforçar as desigualdades existentes, perpetuando a discriminação racial e de gênero no processo de tomada de decisão judicial. Portanto, é fundamental que os vieses algorítmicos sejam reconhecidos, analisados criticamente e abordados para garantir um sistema de justiça mais equitativo e justo para todos os indivíduos envolvidos.

2.5 Possíveis soluções para reduzir ou eliminar os vieses algorítmicos que afetam as decisões de julgamento

A mitigação dos vieses algorítmicos no COMPAS e em outros sistemas similares é uma questão complexa e crucial para garantir a justiça e a equidade nas decisões de julgamento. Diversos autores têm proposto inúmeras soluções e estratégias para reduzir ou eliminar esses vieses, visando alcançar um sistema de justiça mais imparcial e justo.

Uma das possíveis soluções sugeridas é a reformulação das variáveis de entrada do algoritmo. Dressel e Farid (2018) argumentam que a exclusão de informações sensíveis, como raça e gênero, das variáveis utilizadas pelo COMPAS pode ajudar a reduzir os vieses algorítmicos. Eles afirmam que é necessário considerar quais características são verdadeiramente relevantes para a avaliação do

risco de reincidência, buscando evitar a perpetuação de estereótipos e preconceitos no algoritmo.

Outra abordagem proposta é o desenvolvimento e o uso de algoritmos de aprendizado de máquina mais equitativos. Angwin *et al.* (2016) defendem a importância de se considerar a equidade algorítmica no projeto e na implementação de sistemas como o COMPAS. Isso implica em utilizar técnicas que levem em conta a representatividade dos grupos minoritários durante o treinamento do algoritmo, evitando o viés resultante de conjuntos de dados desbalanceados.

A auditoria e a transparência dos algoritmos também são apontadas como medidas essenciais. Barocas e Selbst (2016) destacam que a disponibilização de informações sobre como o algoritmo opera, quais variáveis são consideradas e como os pesos são atribuídos pode ajudar a identificar e mitigar os vieses algorítmicos. A transparência permite uma análise crítica do funcionamento do algoritmo e abre espaço para a detecção de possíveis falhas e injustiças.

Além disso, é fundamental promover uma discussão ampla e inclusiva envolvendo especialistas, a comunidade jurídica e a sociedade civil. Diversos autores ressaltam a importância de abordar a questão dos vieses algorítmicos de forma colaborativa e participativa. Essa discussão coletiva permite uma compreensão mais aprofundada das implicações éticas, sociais e legais dos algoritmos no sistema de justiça e contribui para o desenvolvimento de estratégias mais eficazes no combate aos vieses algorítmicos.

Em suma, para reduzir ou eliminar os vieses algorítmicos nas decisões de julgamento, é necessário reformular as variáveis de entrada, desenvolver algoritmos mais equitativos, promover a auditoria e a transparência dos algoritmos e fomentar uma discussão ampla e inclusiva. Essas medidas visam garantir um sistema de justiça mais imparcial, que trate todos os indivíduos de forma equitativa e justa, independentemente de sua raça, gênero ou outras características pessoais.

3 METODOLOGIA

A metodologia adotada neste estudo envolveu uma pesquisa documental e revisão bibliográfica abrangente, coleta e análise de dados, além de uma revisão crítica da literatura. Essa abordagem permitiu uma análise aprofundada dos vieses algorítmicos no COMPAS, contribuindo para o conhecimento e o debate sobre a justiça criminal e a aplicação de algoritmos nesse contexto.

Inicialmente, realizou-se uma pesquisa documental e revisão bibliográfica, abrangendo estudos e pesquisas relevantes sobre vieses algorítmicos, justiça criminal e sistemas de pontuação de risco. Essa revisão foi fundamental para estabelecer uma base teórica sólida e compreender as principais questões relacionadas ao tema.

Em seguida, para a análise do COMPAS, utilizou-se material livremente disponibilizado pela ProPublica, que consiste em informações sobre indivíduos avaliados pelo algoritmo em questão e as decisões de julgamento subsequentes (ANGWIN *et al.*, 2016).

Foi realizada uma revisão crítica da literatura, buscando a contribuição de autores relevantes que abordaram os vieses algorítmicos e suas implicações no contexto da justiça criminal. Essa revisão permitiu a contextualização dos resultados encontrados e a comparação com estudos anteriores.

Por fim, os resultados obtidos foram cuidadosamente analisados e interpretados à luz da teoria e das discussões acadêmicas existentes. Foi dada ênfase

à discussão das limitações do estudo, bem como às suas implicações para o campo da justiça criminal e a necessidade de medidas corretivas.

É importante ressaltar que este estudo teve como objetivo analisar e compreender os vieses algorítmicos no COMPAS, no entanto, reconhece-se que existem limitações inerentes aos conjuntos de dados utilizados, à metodologia adotada e às próprias ferramentas de pontuação de risco. Portanto, recomenda-se a realização de pesquisas adicionais e a participação de diversos especialistas para aprofundar a compreensão dessas questões e explorar alternativas mais justas e confiáveis.

4 RESULTADOS E DISCUSSÕES

A análise do COMPAS e sua aplicação no sistema de justiça criminal revelam uma série de resultados preocupantes, levantando questões sobre a equidade e a validade do sistema. Estudos baseados em conjuntos de dados e pesquisas empíricas têm contribuído significativamente para a compreensão desses resultados e sua interpretação.

Uma das principais constatações é a presença de vieses algorítmicos no COMPAS. A partir da análise de conjuntos de dados, como o compilado pela ProPublica, evidenciou-se uma tendência do COMPAS em classificar de forma equivocada pessoas negras, atribuindo a elas um risco maior do que o real em comparação com indivíduos brancos (ANGWIN *et al.*, 2016). Esses resultados confirmam a existência de disparidades raciais no sistema de pontuação de risco e levantam preocupações sobre a discriminação algorítmica.

Além disso, estudos mais aprofundados revelam a influência de variáveis questionáveis na geração das pontuações de risco. Dressel e Farid (2018) argumentam que o COMPAS utiliza variáveis, como condenações anteriores, que podem servir como proxies de fatores socioeconômicos e raciais, ampliando os vieses presentes no sistema. Essa análise aponta para a necessidade de revisão e questionamento das variáveis utilizadas no COMPAS, a fim de garantir a validade e a equidade das previsões geradas.

A discussão em torno dos resultados do COMPAS é ampla e complexa, em que diversos autores têm se posicionado a respeito, contribuindo para o debate sobre os vieses algorítmicos e suas implicações. Enquanto alguns concordam com os resultados encontrados, defendendo a busca por alternativas mais justas e transparentes (DIETERICH *et al.*, 2016), outros argumentam que, apesar de suas limitações, o COMPAS ainda pode fornecer informações úteis para apoiar as decisões de julgamento (BERK *et al.*, 2018). Essa discussão reflete a complexidade e as diferentes perspectivas sobre o uso de sistemas de pontuação de risco no contexto da justiça criminal.

Em suma, os resultados e discussões em torno do COMPAS evidenciam a presença de vieses algorítmicos e disparidades no sistema de justiça criminal. A análise crítica desses resultados é essencial para compreender as falhas e desafios enfrentados pelo COMPAS, assim como para buscar soluções mais equitativas e justas. É fundamental um debate amplo e aberto, envolvendo pesquisadores, profissionais do direito e a sociedade como um todo, para avançar na busca por um sistema de justiça mais imparcial e baseado em evidências.

5 CONSIDERAÇÕES FINAIS

A análise dos vieses algorítmicos presentes no COMPAS e sua aplicação no sistema de justiça criminal revelou uma série de desafios e preocupações que merecem atenção. A partir dos conceitos de algoritmo, aprendizado de máquina e vieses algorítmicos, explorou-se o funcionamento do COMPAS, identificando os principais problemas que afetam suas decisões de julgamento.

Mediante a revisão da literatura e a análise de estudos relevantes, foi possível demonstrar que o COMPAS apresenta tendências discriminatórias, especialmente em relação a indivíduos pertencentes a grupos raciais minoritários. Os vieses algorítmicos resultam em disparidades no tratamento e na pontuação de risco, levantando preocupações sobre a equidade e a justiça do sistema de justiça criminal.

Além disso, a análise revelou a influência de variáveis questionáveis no processo de pontuação de risco do COMPAS, o que pode resultar em previsões imprecisas e injustas. É crucial questionar a validade e a relevância dessas variáveis, considerando seus possíveis impactos sociais e raciais.

Diante desses resultados, é imperativo buscar soluções que reduzam ou eliminem os vieses algorítmicos no COMPAS e em outros sistemas de pontuação de risco. A transparência, a auditoria e a responsabilidade devem ser promovidas na implementação e no uso dessas ferramentas. Além disso, é fundamental envolver especialistas multidisciplinares, incluindo cientistas de dados, pesquisadores, profissionais do direito e representantes das comunidades afetadas, para desenvolver abordagens mais justas e equitativas.

Embora seja importante reconhecer os benefícios potenciais do uso de algoritmos no sistema de justiça criminal, é igualmente essencial estar ciente das limitações e dos riscos associados a essas tecnologias. A busca por um sistema de justiça mais justo, imparcial e baseado em evidências requer uma abordagem crítica e colaborativa, que leve em consideração não apenas as capacidades técnicas dos algoritmos, mas também as dimensões éticas, sociais e legais envolvidas.

Em última análise, a questão dos vieses algorítmicos no COMPAS nos desafia a repensar e a aprimorar as práticas de tomada de decisão no sistema de justiça criminal. Devemos buscar um equilíbrio entre o uso de tecnologias avançadas e a garantia dos direitos e da equidade para todos os indivíduos envolvidos. Somente por meio de uma abordagem crítica, colaborativa e informada poderemos avançar em direção a um sistema de justiça mais justo, transparente e confiável.

Como trabalhos futuros, pretende-se replicar o estudo realizado pela ProPublica, utilizando o mesmo conjunto de dados e expandindo-o com entradas mais recentes, a fim de compará-los e investigar a persistência dos vieses algorítmicos no COMPAS. Além disso, busca-se utilizar ferramentas recentes de auditoria de algoritmos para realizar uma análise mais abrangente e precisa dos vieses presentes no sistema.

REFERÊNCIAS

ANGWIN, Julia et al. Machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. **ProPublica**, Nova York, 23 mai. 2016 Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 10 abr. 2023.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **California Law Review**, California, jun. 2016, p. 671-732. Disponível em: <https://doi.org/10.15779/Z38BG31> . Acesso em: 11 abr. 2023.

BERK, Richard et al. Fairness in criminal justice risk assessments: The state of the art. **Sociological Methods & Research**, 2021, v. 50, n. 1, p. 3-44. Disponível em: <https://doi.org/10.1177/0049124118782533> . Acesso em: 11 abr. 2023.

CORMEN, T. H. **Algorithms Unlocked**. 1 ed. MIT Press, 2013.

DIETERICH, William; MENDOZA, Christina; BRENNAN, Tim. COMPAS risk scales: Demonstrating accuracy equity and predictive parity. **Northpointe Inc**, Estados Unidos, 2016, v. 7, n. 7.4, p. 1. Disponível em: [http://go.volarisgraoup.com/rs/430-MBX-989/images/%20ProPublica Commentary Final 070616.pdf](http://go.volarisgraoup.com/rs/430-MBX-989/images/%20ProPublica%20Commentary%20Final%20070616.pdf) . Acesso em: 10 abr. 2023.

DRESSEL, Julia; FARID, Hany. The accuracy, fairness, and limits of predicting recidivism. **Science Advances**, v. 4, n. 1, p. 1-5, jan. 2018. Disponível em: <https://doi.org/10.1126/sciadv.aao5580> . Acesso em: 11 abr. 2023.

HOWARD, A.; BORENSTEIN, J. The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. **Science and engineering ethics**, Springer, v. 24, n. 5, p. 1521–1536, out. 2018. Disponível em: <https://doi.org/10.1007/s11948-017-9975-2> . Acesso em 10 abr. 2023.

MITCHELL, T. **Machine learning**. 1. ed. New York: McGraw Hill, 1997.