

VIÉS ALGORÍTMICO E PODER JUDICIÁRIO: UMA INVESTIGAÇÃO DOS MECANISMOS JURÍDICOS EXISTENTES

 <https://doi.org/10.56238/arev6n4-208>

Data de submissão: 13/11/2024

Data de publicação: 13/12/2024

Fernando José Figueiredo Uchôa de Moura Neto

RESUMO

A presente pesquisa almeja investigar a existência de mecanismos jurídicos eficazes para coibir a influência de ideologia em programas que usam IA na seara judicial. Como objetivos específicos, traz o exame dos fundamentos teóricos do viés algorítmico, a análise da utilização de IA no Poder Judiciário e, por fim, a investigação de possíveis mecanismos de filtragem do viés algorítmico. No que toca metodologia, trata-se de uma pesquisa qualitativa, tendo em vista que a discussão proposta necessita de contextualização e não pode ser reduzida a dados estatísticos. O método hipotético-dedutivo foi utilizado através da formulação da hipótese: “Os mecanismos jurídicos em vigor são suficientes para coibir a influência de ideologias e valores na programação e operação de IA”. Conclui-se que, no caso concreto, os mecanismos jurídicos existentes se mostram insuficientes. A inteligência artificial tem potencial de aumentar a celeridade e eficácia da análise processual, contudo precisa passar por testes rigorosos para afastar possíveis traços discriminatórios. A IA deve ser complementar a tomada de decisão judicial, não um elemento essencial desta.

Palavras-chave: Inteligência Artificial, Viés Algorítmico, Poder Judiciário.

1 INTRODUÇÃO

George Orwell, em sua obra “1984”, imaginou um universo monitorado por máquinas, no qual o livre arbítrio dos seres humanos encontrava-se limitado pela vigilância constante de seus governantes. Não havia a opção de pensar de forma divergente ao governo ou se rebelar a invasão de privacidade promovida por este. As máquinas, nesse contexto, eram programadas para controlar e espionar os homens.

A partir deste exemplo proveniente da literatura, pode-se perceber que a intenção do desenvolvedor ao trabalhar uma certa tecnologia e suas ideologias acabam influenciando a máquina. Assim, a “teletela”, por si só, não reprimia os habitantes de “Ocêania”, na realidade, representava um instrumento governamental programado para tal, reproduzindo os valores e ideologias de seus desenvolvedores.

Nos tempos atuais, a tecnologia avançou de forma ímpar, redes sociais, softwares inteligentes, casas integradas, carros autônomos, são alguns exemplos deste salto tecnológico. O que autores de ficção científica poderiam apenas imaginar no século passado, faz parte do dia a dia do homem século XXI.

A integração da Inteligência Artificial (IA) com a rotina diária é um desses pontos que antes pareciam elementos fantásticos de uma mente criativa. Contudo, é preciso compreender que assim como a “teletela” é fruto da cognição humana, a IA também, logo pode sofrer influências ideológicas daqueles que alimentam sua base de dados.

Os chamados vieses algorítmicos relacionam-se com propensão de um algoritmo em reforçar ou manter padrões discriminatórios presentes na sociedade, o que pode ocorrer considerando que o aprendizado de máquina passa pelo crivo humano tanto na estruturação das bases de dados, quanto no que diz respeito ao conteúdo escolhido para formação desses bancos (SENA, 2023).

Importa destacar que a IA vem ganhando cada vez mais espaço dentro do Poder Judiciário, sendo utilizada para filtrar recursos, criar mecanismos de busca mais eficientes, atender o público através de *chatbots*, dentre outros. Todavia, considerando que a Constituição Brasileira traz como objetivos fundamentais da República Federativa do Brasil a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação, percebe-se a necessidade de buscar a construção de IA que consiga afastar o máximo possível a influência dos valores e ideologias dos desenvolvedores em seu banco de dados.

Assim, como objetivo geral do presente estudo, tem-se o seguinte: perscrutar se existem mecanismos jurídicos eficazes para coibir a influência de ideologias e valores possivelmente ensinados

a IA. Como objetivos específicos: examinar os fundamentos teóricos do viés algorítmico; analisar a utilização de IA no Poder Judiciário; investigar possíveis mecanismos de filtragem do viés algorítmico.

Em relação ao caminho metodológico traçado, tem-se como método de abordagem empregado o qualitativo, considerando que o fenômeno estudado hipotético-dedutivo foi empregado durante a discussão deste trabalho, partindo da ideia de que os algoritmos, por serem programados por humanos, podem ser influenciados pelo viés ideológico destes.

Saliente-se que, o trabalho é fruto de uma revisão bibliográfica sobre a temática em voga. Nesse sentido, a busca por fontes deu-se em sítios online, considerando que a área mater (tecnologia) possui uma forte ligação com publicações nesta espécie de repositórios, como Scielo, Google Acadêmico, Periódicos da CAPES e em repositórios de teses e dissertações acadêmicas. Nos sites citados, realizou-se uma busca por palavras chave “VIÉS ALGORÍTMICO”, “IA+VIÉS”, “PODER JUDICIÁRIO+IA”.

A título de referências primárias foram utilizados artigos que tratam da temática, assim como obras acadêmicas e legislações pertinentes.

2 CONCEITOS PRELIMINARES

A primeira seção do presente estudo versa sobre conceitos caros a pesquisa, notadamente inteligência artificial, algoritmo, big data, banco de dados, todos pertencentes a área da tecnologia com impacto no Poder Judiciário, através de programas empregados na busca pela eficiência dos serviços que presta esta esfera de poder. Busca-se compreender a importância e o impacto das novas tecnologias no Poder Judiciário e a problemática em utilizar inteligência artificial para produção de decisões e sentenças, considerando que o algoritmo pode conter vieses em seu banco de dados.

2.1 A INTELIGÊNCIA ARTIFICIAL E SUAS RAMIFICAÇÕES

A Inteligência Artificial (IA) pode ser definida como uma forma simulação da inteligência humana por sistemas de computador, que realizam atividades que são próprias dos seres humanos, como reconhecimento de fala, tomada de decisões, resolução de demandas, dentre outras. Assim, pode-se afirmar que a IA tem como característica a utilização de uma máquina para o exercício de funções cognitivas (QUEIROZ, DISCONZI, 2024).

Em outra perspectiva, inteligência artificial refere-se à capacidade de máquinas ou sistemas não vivos de resolver problemas e adaptar-se a dificuldades, a fim de alcançar objetivos estabelecidos de forma prévia. Não se configura como essencial que a máquina tenha consciência de sua existência para que consiga desempenhar tais funções, apenas que seja apta a realizar tarefas tipicamente

humanas, como correlacionar informações com a finalidade de apontar incoerências e realizar uma investigação mais aprofundada, dirigir veículos como os carros da Tesla, entre outros (MACHADO SEGUNDO, 2024, p.23).

Nesse sentido, há diferentes tipos e abordagens de IA no universo tecnológico. Atualmente, os sistemas desenvolvidos se voltam a resolução de tarefas específicas e não tem a capacidade de generalização, nem consciência. Todavia, em nível hipotético, busca-se desenvolver uma IA que consiga compreender, entender e aplicar conhecimento de forma geral, tal qual um ser humano (QUEIROZ, DISCONZI, 2024).

Um sistema que utiliza inteligência artificial é alimentado com dados e informações, desta forma vai aprendendo com estes e se adaptando conforme a necessidade apresentada a cada nova entrada de dados. A partir disso, desempenha as funções para as quais é treinado, normalmente atividades que antes eram executadas por pessoas. Note-se que também tem a capacidade de procurar novas formas de solucionar problemas conhecidos a partir dos padrões apresentados nos dados com os quais tem contato (BARBOSA, PORTES, 2019, p.16).

Representa um avanço tecnológico considerável que integra diversos aspectos da ciência e engenharia da computação. Através de sistemas com IA, atividades rotineiras podem ser agilizadas e burocracias desnecessárias são eliminadas, tornando o desempenho de certas tarefas mais eficiente, como a exemplo de mecanismos inteligentes de busca de jurisprudência e o chat GPT, que possui múltiplas funcionalidades, desde tradução e correção de textos, a produção de resumos e ponto-chaves, busca de informações atualizadas, resposta a perguntas dos mais variados assuntos.

Para que uma inteligência artificial possa funcionar, é preciso que haja um banco de dados para que esta aprenda e analise as informações. O *machine learning* é justamente a capacidade da máquina de aprender e tomar decisões de forma autônoma, tendo como parâmetros a reconhecimento de padrões e o tratamento de dados. Através dos mecanismos *deep learning*, que seria um “aprendizado aprofundado”, as IA conseguem aumentar sua capacidade de aprender, empregando verdadeiras redes neurais complexas para este fim (BARBOSA, PORTES, 2019, p.19).

Deste modo, a inteligência artificial se caracteriza como um sistema que mimetiza o pensamento e redes neurais humanas com objetivo de realizar atividades a partir dos bancos de dados aos quais tem acesso.

O acúmulo de grandes quantidades de dados, ou big data, não implica que estes possuam um sentido inerente. Na realidade, os dados necessitam de tratamento, ou seja, seleção, interpretação e análise por parte do algoritmo. Inclusive, faz-se mister compreender que por mais numerosos que

sejam os dados, ainda podem ser falíveis ou representar a realidade de forma imprecisa (MACHADO SEGUNDO, 2024, p.28).

2.2 O ALGORITMO E SUA IMPORTÂNCIA NO CONTEXTO DA IA

Um software lança mão de algoritmos para dar instruções de como algo deve ser feito, quais os componentes necessários, como emprega-los e com a sua finalidade determinada. Algoritmos normalmente estão ligados a operações matemáticas, utilizando equações, aritmética, álgebra, cálculo, lógica e probabilidade, traduzidas em forma de código (MACHADO SEGUNDO, 2024, p.24).

Nesse sentido, Carvalho (2020, p.27) afirma que:

No contexto da ciência da computação, os programadores (humanos) são os profissionais responsáveis pelo desenho de fluxos lógicos nos quais, a partir de dados de entrada determinados (informações inseridas pelo usuário do sistema, como por exemplo respostas de questionários dinâmicos ou ainda interpretação de texto feita por algoritmos), seja possível obter uma série de dados de saída desejados (o resultado da operação, informação condicionada e processada de acordo com os dados inicialmente inseridos na operação pelo usuário do sistema). Assim, os programadores fazem uso de paradigmas relacionados à construção do raciocínio que deverá ser utilizado pelo algoritmo, que, por sua vez, será estruturado de modo a resolver um problema objetivo.

Ainda, afirma Machado Segundo (2024, p 24) o algoritmo é derivado de um modelo, um fragmento do mundo que o rodeia, do qual recebem informações e dados com a finalidade de interpretá-los. Saliente-se que nem sempre a representação feita por um algoritmo é fiel a realidade, haja vista que podem existir interferências da visão e valores daqueles que alimentam o sistema com os dados.

Assim, os algoritmos consistem em modelos matemáticos desenvolvidos para determinado fim e que podem ser limitados por opiniões ou vieses contidos no código. Deste modo, o homem ensina a máquina a pensar e nesse processo pode embutir nela dados enviesados, que corroborem com sua visão de mundo de forma específica (VIEIRA, 2019).

Neste sentido, Machado Segundo (2024, p.24) ensina que:

“(...) os algoritmos perseguem objetivos, os quais, até o momento, lhes são externos, ou seja, são indicados por quem os elabora ou idealiza. Dessa forma, a discussão sobre quais objetivos os algoritmos devem perseguir antecede e transcende qualquer debate a respeito de sua aceitabilidade ou utilidade, evidenciando ainda o fato de que eles não são neutros. Como advertem Maria do Céu Patrão Neves e Maria da Graça Carvalho, toda tecnologia é obra do ser humano e, como tal, “traz em si, irredutivelmente inculcada, a marca do seu criador que se expressa sob o signo da finalidade” (MACHADO SEGUNDO, 2024, p.24).

Destaque-se que não são propriamente softwares, na realidade, se caracterizam como procedimentos em código, que empregam determinados cálculos, convertendo dados em soluções

específicas. Por exemplo, um algoritmo que se adapta a mudança de rota de um ponto A ao ponto B em questão de segundos (GILLESPIE, 2018, p. 97).

Ainda, importa Gillespie (2018, p. 98) ensina que os algoritmos possuem seis dimensões de relevância pública com valor político: padrões de inclusão, ciclos de antecipação, avaliação de relevância, promessa de objetividade algorítmica, entrelaçamento com a prática, produção de públicos calculados.

Os padrões de inclusão se relacionam com a necessidade de bancos de dados para que os algoritmos ganhem sentido. Políticas de coleta de dados que observem requisitos mínimos de confiabilidade são fundamentais neste processo, afinal, os bancos de dados nada mais são do que a junção de um conjunto previamente coletado de informações que serve um determinado fim (GILLESPIE, 2018, p.101).

Além disso, muitos algoritmos são trabalhados com a finalidade de antecipação de comportamento dos usuários, usando termos estatísticos e dados demográficos para cruzar informações e engajar os usuários (GILLESPIE, 2018, p.101).

Nesse sentido, o algoritmo, igualmente, realiza uma avaliação de relevância através de indicadores, que empregam cálculos complexos contendo dados contextuais e técnicas de processamento de linguagem natural a fim de obter melhores resultados no que toca a relevância de um conteúdo para um determinado usuário (GILLESPIE, 2018, p.102).

Por fim, saliente-se que os algoritmos configuram estabilizadores de confiança, supostamente livres de quaisquer subjetividades, produtores de avaliações rigorosas e corretas, alheios a erros e tentativas de influência. Todavia, não há como existir isenção total em um serviço de buscas. Por exemplo, se um usuário procurar algo que fere algum indicador previamente estabelecido com a finalidade de proteger determinado grupo, como conteúdos impróprios com menores, não obterá resultados em buscadores tradicionais como Google (GILLESPIE, 2018, p.107).

Assim, nota-se que a imparcialidade de um algoritmo pode ser modulada conforme as suas finalidades.

2.2.1 Vieses Algorítmicos: até onde vai a imparcialidade dos bancos de dados?

A inteligência de uma IA é fruto da cognição de um ser humano, assim, preconceitos, valores, ideologias podem tornar seu processo de tomada de decisão ou resolução de problemas falho (VIEIRA, 2019).

Cabe destacar que a informação pode ser vislumbrada como símbolo frutos da cognição humana, a exemplo de palavras orais ou escritas. Igualmente, representar objetos, movimentações, tudo depende do contexto em que se insere.

Nos temos atuais, ganhou contornos diferenciados, se tornando primordial. Quem a detém, dentro da lógica capitalista, consegue uma vantagem, visto que se baseia na comunicação e conexão. Fundamenta-se na ideia de transparência, ou seja, a imposição por parte do sistema de tornar tudo disponível como informação. Assim, as informações circulam de forma livre, mas as pessoas não (HAN, 2022, p.8).

Neste sentido, um algoritmo é impactado pela qualidade dos dados de treinamento e também pela variedade e quantidade de interações com usuários de verdade. Ainda, faz-se uso de Modelos de Linguagem Grande, para obtenção de respostas e tomadas de decisão que tenham relevância dentro do contexto no qual se inserem (OLIVEIRA, ARANTES, 2024, p. 7).

Deste modo, percebe-se que os dados podem ser manipulados por aqueles que a detém para cumprir suas próprias agendas, o que pode macular até mesmo código. A ideologia dos coletores dos dados pode afetar a forma que os parâmetros são definidos, produzindo “vieses algorítmicos”.

Saliente-se que os seres humanos estão cada vez mais integrados com a IA, de forma que esta influência em suas decisões diárias. Seja a partir de uma sugestão de local recomendado conforme seu perfil de busca ou uma propaganda adaptada as interações dentro de rede social, os usuários têm suas vontades atravessadas pelos dados aos quais são expostos (OLIVEIRA, ARANTES, 2024, p. 6).

Nesta toada, Oliveira e Arantes (2024, p.7) defendem que:

Os problemas relativos à suposta neutralidade do algoritmo não são novos e estão presentes no uso da internet há algum tempo. Sobre os mecanismos de busca e pesquisa, por exemplo, a pesquisadora Safiya Umoja No-ble (2013) considera que **a implementação desses algoritmos nem sempre é neutra, pois eles determinam, a partir de diversos fatores não muito claros aos usuários, quais assuntos e conteúdos terão maior ou menor relevância, aderência e potencial de engajamento do público**. Desse modo, é possível afirmar que muitas vezes os algoritmos incorrem no risco de perpetuar e amplificar, mesmo que de maneira sutil, as desigualdades sociais e raciais existentes (grifo próprio).

Deste modo, nota-se que o viés algorítmico deve ser combatido, embora se pense que computadores sejam neutros por excelência, os programadores e os próprios bancos de dados podem ser afetados por ideologias e preconceitos. Assim, os resultados apresentados não podem ser considerados totalmente imparciais.

Nesta toada, os bancos de dados podem funcionar como uma espécie de viés algoritmo, haja vista que um determinado filtro de dados pode partir de um preconceito ou visão particular de quem

estabeleceu que aquele seria considerado um indicador válido para composição de um grupo de dados específico.

Saliente-se que não uma solução simples para esta questão, afinal, os bancos de dados são necessários para o bom funcionamento dos programas que usam inteligência artificial. Igualmente, deve-se pensar que por mais que possa representar a estatística uma espécie de viés, sua importância no contexto computacional e tecnológico é extrema. Assim, faz-se mister encontrar um equilíbrio entre os indicadores selecionados para compôr um banco de dados, evitando concentrar filtros apenas em grupos específicos em detrimento de outros.

Quando há desbalanceamento nos dados de treinamento de uma inteligência artificial, os algoritmos acabam reforçando preconceito, discriminação e geram desigualdades ao invés de combate-las.

Obviamente, não é possível um algoritmo agir de forma autônoma, afinal, como visto anteriormente, é composto por um conjunto de instruções programadas com um determinado fim. Desta maneira, o rastreamento de possíveis vieses algoritmos é relevante para que o uso da tecnologia seja feito de forma consciente e conforme os preceitos constitucionais.

2.2.2 Viés Algorítmico na prática: algoritmos podem reproduzir preconceitos e discriminar?

Nesta seção, pretende-se esmiuçar alguns casos comprovados de viés algoritmo tanto no contexto de discriminação de gênero, quanto racial, objetivando responder à pergunta chave proposta: algoritmos podem reproduzir preconceitos e discriminar?

Primeiramente, salienta-se que, para fins de estudo, foram selecionados dois aspectos do viés algoritmo: a discriminação de gênero e a discriminação racial feita por inteligência artificial. A discussão proposta parte de alguns casos práticos comprovados de viés algoritmo.

No que toca a discriminação de gênero, esta caracteriza uma dinâmica que afeta a sociedade como um todo, considerando que há um notável desequilíbrio social entre homens e mulheres. Pesquisas produzidas pela Organização Internacional do Trabalho (OIT), apontam que as mulheres recebem aproximadamente 77% dos salários dos homens. Caso não haja intervenção nesta seara, a igualdade salarial entre gêneros só será obtida em 2086 (SIQUEIRA, 2017, p.289).

Saliente-se que a diferença entre sexo e gênero reside no fato que o primeiro é ligado a questões meramente biológicas, ao passo que o segundo compõe a identidade de um ser humano, sendo influenciado pela cultura e instituições sociais. (SIQUEIRA, 2017, p.291).

Note-se que gênero pode ser vislumbrado como uma categorização pessoal e social de indivíduos, que alicerça a construção de sua identidade, podendo não ser necessariamente compatível com o sexo biológico (SIQUEIRA, 2017, p.289).

Percebe-se até os dias atuais a noção patriarcal de que as mulheres devem ser dominadas por homens devido ao seu gênero. Na antiguidade oriental, a mulher era vista como propriedade masculina, não sendo sujeita quaisquer espécies de direitos. Igualmente, no período da Idade Média, a mulher, consoante os ditamente do cristianismo, deveria se reportar ao seu pai e após o seu casamento, ao marido (SIQUEIRA, 2017, p.294).

Com a idade moderna, houve uma mudança de paradigma, as mulheres passaram a ter uma posição mais ativa, exigindo liberdade e autonomia. Começam a ganhar espaço no mercado de trabalho, contudo a discrepância salarial é notável entre homens e mulheres, mesmo quando exercendo funções iguais (SIQUEIRA, 2017, p.294).

Assim, percebe-se que a discriminação de gênero tem uma raiz histórica e atravessa civilizações. Atualmente, há inúmeras legislações que buscam atuar no desbalanço entre homens e mulheres, a fim de garantir igualdade material entre estes. Além disso, a discussão e combate a discriminação de gênero pode ser observada em diversos setores sociais.

Neste sentido, homens e mulheres, consoante o artigo quinto inciso primeiro da Constituição Federal, são iguais em direitos e obrigações. Todavia, a diferença de tratamento entre os gêneros pode ser visualizada em diversas vertentes, inclusive no universo digital.

Um notório exemplo dessa desigualdade envolve o Apple Card e seu algoritmo. Percebeu-se, em novembro de 2019, que quando o Apple Card designava limites de crédito, as mulheres obtinham uma quantidade inferior aos cônjuges, apesar de partilharem da mesma renda e pontuação de crédito (SENA, 2023).

Em decorrência deste viés, a Apple foi investigada, três meses após o lançamento do serviço, por órgãos regulatórios dos Estados Unidos, considerando o sexismo em sua concessão de crédito. O caso ganhou notoriedade após um programador da Dinamarca, Heinemeir Hansson, relatar em suas redes sociais que sua esposa, mesmo sendo melhor remunerada, teve um limite designado vinte vezes menor que o seu (AMARAL, 2020, p.13).

Curiosamente, a esposa do co-fundador da empresa que trabalhou no desenvolvimento do cartão de crédito, Steve Wozniak, igualmente teve um limite designado dez vezes menor que o deu seu marido (AMARAL, 2020, p.13).

Outro notável exemplo da discriminação de gênero feita por inteligência artificial envolve o sistema de tradução automatizada google tradutor. Ao colocar certas frases em inglês para tradução

em português como “i want to be a judge”, a inteligência artificial retorna como resultado “eu quero ser UM JUIZ”. Ao passo que a frase “i want to be a nurse” é traduzida como “eu quero ser UMA ENFERMEIRA” (MACHADO SEGUNDO, 2024, p.38).

Igualmente, observa-se este viés a partir da tradução da seguinte frase “the judge and her husband”, que tem como resultado “o juiz e o marido dela”. Percebe-se que a posição de juiz, para o algoritmo, é ocupada por homem, já que a inteligência artificial não faz sequer uma leitura contextual da frase, que levaria a tradução “a juíza e seu marido” (MACHADO SEGUNDO, 2024, p.41).

Outro teste que demonstra perfeitamente a questão da discriminação de gênero em relação ao google tradutor é a frase “i am a woman and i want to be a judge”. Nesse caso, o tradutor retorna “eu sou uma mulher e quero ser um juiz”. Novamente, o gênero não é analisado de forma contextual, a palavra juiz é traduzida na versão masculina mesmo que a frase seja construída a partir da perspectiva feminina (MACHADO SEGUNDO, 2024, p.41).

Outrossim, outro notório caso envolvendo viés algorítmico e questões de discriminação de gênero se refere ao sistema desenvolvido pela Amazon para análise de currículos de candidatos a postos de trabalho na empresa. A inteligência artificial foi programada para revisar os currículos e atribuir uma pontuação a estes que variava entre uma e cinco estrelas. Notou-se, após algum tempo de utilização, que candidatas do sexo feminino não eram classificadas para cargos técnicos (SALMORIA, 2024, p.5).

Após uma investigação para descobrir as causas desse viés, os responsáveis pelo projeto compreenderam que o banco de dados empregado para treinar a IA tomou por base currículos enviados a Amazon nos dez anos anteriores e a maior parte das contratações efetivadas referentes a estes foi de candidatos do sexo masculino. Por tal razão, a IA entendeu que era preferível pontuar de forma mais significativa currículos de homens em detrimento de mulheres (SALMORIA, 2024, p.6).

No que toca a questão de discriminação racial, o viés algorítmico igualmente já foi constatado em alguns programas que utilizam IA dentro do contexto judicial e policial.

Há relatos de racismo no software de reconhecimento facial utilizado pela polícia dos Estados Unidos. Conforme a ACLU (União da Liberdade Civil Americana), um homem chamado Robert Williams foi detido indevidamente, por aproximadamente 30 horas, devido a um erro do sistema de reconhecimento facial da Polícia do Estado de Michigan. A partir desse erro, notou-se que o software adotado não realizava a devida distinção entre pessoas negras, considerando que as únicas características em comum entre Robert e o suspeito eram a cor de sua pele e altura (SENA, 2023).

Saliente-se que a Constituição Federal Brasileira elegeu como objetivo fundamental da República Brasileira a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação.

Ainda, pode-se apontar como caso emblemático de viés algorítmico o sistema COMPAS (Perfil de Gerenciamento Corretivo de Infratores para Sanções Alternativas), elaborado pela empresa Northpointe, que visava realizar a avaliações de risco acerca da probabilidade de reincidência no sistema criminal, além de auxiliar o judiciário norte-americano em informações de decisões e mitigação dos riscos futuros. A partir de um estudo produzido por um jornal investigativo, percebeu-se que o programa em questão era racialmente enviesado (VIEIRA, 2019).

Com os dados de pontuação de risco obtidos de mais de 7 mil pessoas presas em condado da Flórida, entre os anos de 2013 e 2014, percebeu-se que o score de avaliação de risco apresentado pela empresa indicava que pessoas negras seriam de alto risco, ao passo que pessoas brancas eram classificadas como de baixo risco (VIEIRA, 2019).

Insta destacar que a cor não se apresentava como uma variável avaliada de forma explícita no algoritmo, contudo, raça e gênero eram parte integrante de outras variáveis, o que influenciava na classificação de risco (VIEIRA, 2019).

O caso COMPAS demonstra, igualmente, o perigo na utilização de softwares preditivos na atividade judicial, considerando que estes atuam de forma contrária aos ditames da presunção de inocência, ampla defesa, contraditório, haja vista que a pessoa é punida de forma preventiva por atos que poderia cometer no futuro, sem que seus direitos e garantias fundamentais sejam observados.

Oliveira e Arantes (2024, p. 8) afirmam que:

Os problemas relativos à suposta neutralidade do algoritmo não são novos e estão presentes no uso da internet há algum tempo. Sobre os mecanismos de busca e pesquisa, por exemplo, a pesquisadora Safiya Umoja No-ble (2013) considera que a implementação desses algoritmos nem sempre é neutra, pois eles determinam, a partir de diversos fatores não muito claros aos usuários, quais assuntos e conteúdos terão maior ou menor relevância, aderência e potencial de engajamento do público. Desse modo, é possível afirmar que muitas vezes os algoritmos incorrem no risco de perpetuar e amplificar, mesmo que de maneira sutil, as desigualdades sociais e raciais existentes.

Assim, percebe-se a necessidade de combater o viés algorítmico, tornando-a IA uma ferramenta segura para utilização dentro do ambiente judicial.

A redução de um ser humano a um dado estatístico, por si só, acaba revelando um viés, um momento perpetuado no tempo que pode ou não refletir o futuro de uma pessoa. Os seres humanos e sua rotina tendem a ficar cada vez mais integrados com a IA, contudo, esta não pode substituir completamente a função cognitiva de um juiz, que vai muito além dos fatos em si, considerando as

circunstâncias, a personalidade, a conduta do réu, o comportamento da vítima, a culpabilidade, os antecedentes, os motivos.

2.2.3 O viés algoritmo e o Poder Judiciário: há como utilizar IA de forma segura no contexto judicial?

Primeiramente, é preciso compreender que o Poder Judiciário vive uma constante busca por aperfeiçoamento e eficiência na prestação judicial. Deste modo, ferramentas que estimulem a celeridade dos atos processuais são importantes para alcançar uma maior efetividade e eficácia dos serviços prestados por esta esfera do poder, garantindo a preservação da segurança jurídica.

Nesse sentido, a IA tem potencial de auxiliar o Poder Judiciário a alcançar uma maior agilidade na análise processual, todavia, o seu uso para atividades que sejam relativas à tomada de decisão deve passar por rigorosos testes, a fim de mitigar o viés algorítmico. Em relação a atividade ordinárias, a automação via IA já vem sendo observada em diversas instâncias do Poder Judiciário, como a exemplo do projeto VitóriaIA do Supremo Tribunal Federal (ROSSETTI, 2024).

O projeto VitóriaIA foi desenvolvido com objetivo de identificar, no acervo de processos do STF, aqueles que versam sobre os mesmos assuntos e agrupá-los de forma automática. Desta maneira, processos que podem ser trabalhados conjuntamente ou resultar em novos temas de repercussão geral são mais facilmente identificados.

Atualmente, a IA pode ser percebida em diversas áreas do Poder Judiciário, desde a distribuição da petição inicial a execução de algumas atividades próprias de cartórios judiciais (BESSA, 2024, p.8)

Além disso, a IA é amplamente empregada com a finalidade de analisar grandes volumes de dados processuais, possibilitando, assim, uma maior celeridade na tramitação dos processos e auxiliando a tomada de decisões judiciais. Em relação a esta última, é preciso compreender que a IA pode apontar padrões e tendências que ajudam o magistrado em processo decisório. Todavia, não deve ser o fundamento principal para uma decisão judicial, afinal, esta se baseia na cognição do juiz, não da máquina (BESSA, 2024, p.11).

Assim, um dos desafios na utilização de IA dentro do Poder Judiciário consiste em buscar equilíbrio entre cognição humana e máquina, haja vista que a tomada de decisão não envolve apenas aspectos puramente objetivos, os quais a máquina tem uma precisão maior em aferir. Na realidade, o processo judicial perpassa conhecimentos diversos e envolve questões que não podem ser reduzidas a dados estatísticos.

Nesse sentido, Bessa (2024, p.12) ensina que:

As críticas ao positivismo tecnológico apontam para a perda da humanidade nas decisões judiciais. A automação excessiva pode levar à padronização e uniformização das decisões, desconsiderando as peculiaridades de cada caso concreto. A análise puramente técnica pode negligenciar aspectos subjetivos e contextuais que são essenciais para uma justiça individualizada e sensível às particularidades dos envolvidos. Além disso, a falta de compreensão dos algoritmos utilizados nos sistemas inteligentes pode gerar desconfiança e questionamentos sobre a legitimidade das decisões automatizada.

O Poder Judiciário deve manter o fator humano das decisões judiciais preservados, até para evitar alicerçar suas decisões em programas sujeitos a vieses de gênero e raciais, como os apresentados na seção anterior. Elementos subjetivos e contextuais necessários para a produção de uma sentença adequada não conseguem ser devidamente processados por máquinas, apenas a cognição humana é apta a atuar efetivamente na produção de justiça individualizada e pertinente as especificidades dos envolvidos (BESSA, 2024, p. 12).

A tecnologia almeja neutralidade, porém esta nem sempre é alcançada de fato, haja vista os bancos de dados podem ser alimentados a partir de pré-concepções dos sujeitos que trabalham na coleta de dados. Além disso, os dados refletem o meio social que é extremamente afetado por relações de desigualdade, discriminações, exclusões. Logo, a pretensão de desenvolver IA neutra no contexto judicial não se aplica, os próprios dados não o são.

3 MECANISMOS DE FILTRAGEM DO VIÉS ALGORÍTMICO: HÁ COMO PRESERVAR A SEGURANÇA JURÍDICA EM SUA UTILIZAÇÃO?

Insta salientar que, em caso de processos repetitivos e de baixa complexidade, a IA, quando devidamente acompanhada por um juiz que seja responsável pela decisão, pode ser empregada objetivando a celeridade e economia na prestação judicial (BESSA, 2024, p.17).

Todavia, esta supervisão deve ser realizada com alguns cuidados, para evitar que o juiz torne sua cognição dependente da concordância do resultado apontado pela IA. Nesses termos, Bessa (2024) afirma que:

(...) problemas associados a uma excessiva dependência dessa supervisão. Uma das razões para esse problema é o chamado “viés de automação” (ou “machine bias”), que é outro viés cognitivo humano. Esse viés envolve a tendência natural dos seres humanos de favorecer os resultados gerados por sistemas automatizados, devido à crença de que esses resultados são fundamentados em cálculos matemáticos e, portanto, são considerados científicos e precisos. Isso pode levar o tomador de decisão a reduzir sua discordância em relação aos resultados dos sistemas de IA, aceitando-os integralmente ou parcialmente, seja porque não reconhece quando os sistemas automatizados cometem erros, seja porque não dá importância a informações que possam contradizê-los (BESSA, 2024, p.17).

Percebe-se que a utilização de inteligência artificial pode ser uma ferramenta importante, mas não o fundamento último de uma decisão judicial, considerando que não substitui as nuances da compreensão humana e está sujeita a viés algorítmico.

A fim de diminuir o impacto de vieses algorítmicos, faz-se necessário um trabalho colaborativo entre desenvolvedores, pesquisadores e reguladores. Buscar mecanismos de seleção e limpeza de dados de treinamento que consigam filtrar possíveis vieses, além do acompanhamento através de auditorias e testes de forma constante para apurar se há algum desequilíbrio a ser corrigido (OLIVEIRA, ARANTES, 2024, p. 14).

Outrossim, devem ser criadas regulamentações e diretrizes que possam asseverar que os algoritmos sejam imparciais e inteligíveis. A responsabilização na ocorrência de um viés também deve ser observada para certificar que não somente haja a correção do viés, como a punição daquele que o criou (OLIVEIRA, ARANTES, 2024, p. 14).

A participação democrática dos sujeitos de direito também se revela essencial para a correta utilização da inteligência artificial, seja para definir regras e limites para IA no processo civil, como também fomentar debates sobre novas tecnologias no meio social e mitigação de riscos.

4 CONSIDERAÇÕES FINAIS

A inteligência artificial (IA) tem um potencial considerável para transformar o meio social, inclusive a seara judicial. Porém, salienta-se que, mesmo que IA consiga assumir funções cognitivas complexas, não se exime de erros, principalmente no que diz respeito ao viés algorítmico. Os exemplos citados como Apple Card e o sistema de reconhecimento facial utilizado pela polícia do Estado de Michigan, demonstram que os algoritmos podem reproduzir padrões discriminatórios e manter desigualdades sociais, atuando como um reflexo de preconceitos inseridos em seu processo de desenvolvimento.

Outrossim, como pode ser vislumbrado a partir do caso do algoritmo COMPAS, a utilização de softwares preditivos pode acarretar em erros judiciários graves. Além disso, o uso dessas tecnologias pode ir de encontro com princípios e garantias fundamentais como a presunção de inocência e o direito à ampla defesa. A punição prévia dos indivíduos com base em uma previsão enviesada contrária a Constituição Federal.

A partir destas considerações, é possível perceber que a urgência em combater o viés algorítmico e assegurar que IA seja uma ferramenta segura e justa. Faz-se mister a realização de estudos e testes rigorosos dos softwares que utilizem IA na seara jurídica, para identificar e afastar possíveis vieses algorítmicos presentes. Quanto a hipótese levantada, nota-se que a resposta é negativa,

haja vista que os mecanismos jurídicos existentes ainda são insuficientes para afastar o viés algorítmico, porém essa realidade deve ser enfrentada com seriedade e diligência.

Finalmente, o Poder Judiciário pode ganhar muito com a integração a IA, principalmente nos aspectos da celeridade processual e eficiência. Contudo, essa tecnologia deve funcionar como complementar, especialmente no campo de tomada de decisões judiciais. Cautela e responsabilidade são essenciais nesse processo, a fim de buscar o melhor uso da IA afastando-se de possíveis aspectos discriminatórios implícitos no desenvolvimento desta.

REFERÊNCIAS

AMARAL, Camila Augusta Medeiros Colás. Os desafios do uso de algoritmos de inteligência artificial em processos decisórios à luz do ordenamento jurídico. Trabalho de Conclusão de Curso – Escola de Direito, Universidade do Estado do Amazonas. Manaus, 2021.

BESSA, Aline Sousa. Inteligência Artificial E Processo Civil: Riscos Do Positivismo Tecnológico No Decisionismo Judicial. Temas de Direito e Processo, p. 7, 2024.

CARVALHO, Allan Pereira de. Viés algorítmico e discriminação: possíveis soluções regulatórias para o Brasil. Trabalho de Conclusão de Curso – Centro de Ciências Jurídicas e Sociais, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2020.

GILLESPIE, Tarleton. A relevância dos algoritmos. Parágrafo, v. 6, n. 1, p. 95-121, 2018.

HAN, Byung-Chul. Infocracia: digitalização e a crise da democracia; tradução de Gabriel S. Philipson. – Petrópolis, RJ: Vozes, 2022.

MACHADO SEGUNDO, Hugo de Brito. Direito e Inteligência Artificial: o que os algoritmos têm a ensinar sobre interpretação, valores e justiça. 2. ed. - Indaiatuba, SP: Editora Foco, 2024.

OLIVEIRA, J. C. de; ARANTES, P. A. C. Alexa, are you racist? Algorithmic racism, biases and intentionality. DAT Journal, [S. l.], v. 9, n. 1, p. 04–16, 2024. DOI: 10.29147/datjournal.v9i1.749. Disponível em: <https://datjournal.emnuvens.com.br/dat/article/view/749>. Acesso em: 29 sep. 2024.

QUEIROZ, Gabriel Noll; DISCONZI, Verônica Silva do Prado. O IMPACTO DA INTELIGÊNCIA ARTIFICIAL NO DIREITO: QUESTÕES ÉTICAS E LEGAIS. Revista Ibero-Americana de Humanidades, Ciências e Educação, v. 10, n. 4, p. 1388-1406, 2024.

ROSSETTI, Regina et al. Direitos fundamentais no uso de inteligência artificial no poder judiciário brasileiro. Revista Tecnologia e Sociedade, v. 20, n. 59, p. 219-235, 2024.

SALMORIA, C.; ALLESSI NICOLETTI ALVES, D.; ANTUNES DE OLIVEIRA, M. Viés Algorítmico De Gênero No Sistema De Recrutamento:O Caso Amazon. Cadernos UNDB – Estudos Jurídicos Interdisciplinares, [S. l.], v. 7, n. 1, 2024. Disponível em: <https://periodicos.undb.edu.br/index.php/cadernosundb/article/view/199>. Acesso em: 29 set. 2024.

SENA, Erick Morais de. Viés na IA: como o viés algoritmo influencia na perpetuação de estereótipos e desigualdades existentes. Trabalho de Conclusão de Curso – Centro de Engenharia Elétrica e Informática, Universidade Federal de Campina Grande. Campina Grande, 2023.

VIEIRA, Leonardo Marques. A problemática da inteligência artificial e dos vieses algorítmicos: caso COMPAS. Brazilian Technology Symposium. 2019.