

Inteligência Artificial e Aprendizado de Máquina na sociedade moderna

Hermenegildo Woropo Albino Paiva

Mestre em Ciências da Computação

Instituição: Universidade Federal de Lavras

Endereço: Minas Gerais, Brasil

E-mail: hermen.velho02@gmail.com / hermenegildo.paiva@estudante.ufla.br

RESUMO

Nas últimas décadas, a Inteligência Artificial (IA) e o Aprendizado de Máquina (AM) deixaram de ser meramente conceitos da área de ciência da computação, tornando-se fundamentais na sociedade atual. Elas estão presentes em assistentes artificiais, diagnósticos na área da saúde, sistemas de segurança e processos de automação industrial. Entretanto, essa rápida expansão da IA também levanta novos desafios éticos, sociais e regulatórios. Com algoritmos assumindo papéis antes únicos dos humanos, aparecem preocupações sobre a proteção dos dados, a responsabilidade por decisões automatizadas, os preconceitos nos sistemas e a potencial substituição da força de trabalho humana. Assim, este trabalho tem como objetivo analisar a utilização da Inteligência Artificial e do Aprendizado de Máquina na sociedade atual, com foco nas áreas de saúde, segurança e automação. Conclui-se que, a habilidade de analisar grandes quantidades de dados rapidamente e identificar padrões úteis para a tomada de decisões tem permitido progresso que antes parecia impossível, aprimorando diagnósticos médicos, aumentando a vigilância em tempo real e automatizando processos industriais de maneira mais eficiente. No entanto, essas vantagens não vêm sem questões éticas, políticas e legais. O potencial da IA, apesar de promissor, também apresenta vários desafios que demandam análise cuidadosa e regulação atenta.

Palavras-chave: Inteligência Artificial. Aprendizado de Máquina. Ética em IA. Automação. IA na saúde. IA na segurança pública.

1 INTRODUÇÃO

Nas últimas décadas, a Inteligência Artificial (IA) e o Aprendizado de Máquina (AM) deixaram de ser meramente conceitos da área de ciência da computação, tornando-se fundamentais na sociedade atual. Elas estão presentes em assistentes artificiais, diagnósticos na área da saúde, sistemas de segurança e processos de automação industrial, provocando uma mudança significativa na maneira como as pessoas interagem com a tecnologia, fazem escolhas e executam atividades complexas. O desenvolvimento de redes neurais, do processamento de linguagem natural e da análise preditiva está revolucionando setores inteiros, trazendo promessas de maior eficiência, personalização e agilidade.

Entretanto, essa rápida expansão da IA também levanta novos desafios éticos, sociais e regulatórios. Com algoritmos assumindo papéis antes únicos dos humanos, aparecem preocupações sobre a proteção dos dados, a responsabilidade por decisões automatizadas, os preconceitos nos sistemas e a potencial substituição da força de trabalho humana. Por isso, é fundamental entender as diferentes dimensões do uso da inteligência artificial na atualidade, a fim de assegurar que seu progresso respeite os princípios de justiça,



segurança e equidade.

A escolha de examinar a aplicação da Inteligência Artificial e do Aprendizado de Máquina na sociedade contemporânea se justifica pelo papel cada vez mais central que essas tecnologias desempenham em diversas áreas da vida. O impacto já evidente em setores como saúde, segurança pública e automação demonstra que os sistemas inteligentes não são apenas ferramentas, mas transformadores de realidades sociais. Pensar de forma crítica sobre essas mudanças é fundamental para garantir que os benefícios da IA sejam amplamente compartilhados e que seus riscos sejam controlados por meio de uma governança ética e responsável.

Apesar das inovações tecnológicas, a sociedade ainda tem uma falta de diretrizes claras para enfrentar os problemas éticos, sociais e legais que surgem com a utilização da IA. Muitos setores implementam sistemas inteligentes sem levar em conta totalmente suas repercussões sobre os direitos humanos, a equidade e a transparência. Nesse cenário, surge a seguinte preocupação: quais são as principais utilizações, desafios e implicações éticas ligadas ao uso da inteligência artificial e do aprendizado de máquina nas áreas da saúde, segurança e automação, e como essas questões podem ser abordadas de maneira crítica e responsável?

Assim, este trabalho tem como objetivo geral analisar a utilização da Inteligência Artificial e do Aprendizado de Máquina na sociedade atual, com foco nas áreas de saúde, segurança e automação. Especificamente, visa identificar as principais aplicações dessas tecnologias, debater os desafios enfrentados por seus usuários e desenvolvedores, e refletir sobre os dilemas éticos que surgem de sua adoção em larga escala.

O tipo de pesquisa utilizada é qualitativa, centrada em uma revisão de literatura. De acordo com Gil (2008), a revisão bibliográfica é realizada utilizando material que já foi publicado, incluindo livros, artigos acadêmicos, conteúdos digitais e relatórios técnicos.

2 DESENVOLVIMENTO

2.1 FUNDAMENTOS DA INTELIGÊNCIA ARTIFICIAL E DO APRENDIZADO DE MÁQUINA

Segundo Ludermir (2021), a inteligência artificial é uma área da computação dedicada à criação de sistemas que podem emular funções da inteligência humana, como raciocínio, aprendizado, percepção, decisão e entendimento de linguagem natural. Embora essa ideia tenha surgido na década de 1950 com figuras como Alan Turing e John McCarthy, somente nos últimos tempos os progressos na computação e a grande quantidade de dados disponíveis possibilitaram que a IA se tornasse uma das inovações mais impactantes do mundo digital. A IA pode ser dividida em duas principais categorias: a IA fraca, que se concentra em tarefas específicas (como chatbots ou sistemas de recomendação), e a IA forte, que ainda está



em fase de testes, almejando replicar uma inteligência geral e autônoma semelhante à humana.

O aprendizado de máquina, ou Machine Learning (ML), é uma subdivisão da IA que se dedica a técnicas e algoritmos que possibilitam que os sistemas desencadeiem aprendizado a partir de dados, modificando seu desempenho sem a necessidade de intervenção humana conforme aponta Bigonha (2025). Em vez de serem programados com diretrizes fixas, esses sistemas recebem grandes quantidades de dados e, a partir deles, reconhecem padrões e realizam previsões ou decisões. Esse aprendizado pode ocorrer de maneira supervisionada, onde o sistema é treinado com exemplos rotulados (como raios-x identificados como saudáveis ou doentes), ou não supervisionada, onde o sistema tenta descobrir agrupamentos e formas dentro dos dados sem orientações claras. Existe também o aprendizado por reforço, em que a máquina adquire conhecimento por meio de tentativas e erros, com base em recompensas e punições.

É relevante ressaltar a diferença entre IA e ML, embora ambas se complementem: enquanto a IA busca criar máquinas com inteligência, o ML serve como uma das maneiras mais eficientes de alcançar esse objetivo. Outras áreas dentro da IA incluem o processamento de linguagem natural, a visão computacional, sistemas peritos e robótica autônoma. O progresso nesses campos é principalmente influenciado pela qualidade dos dados, pela capacidade de processamento computacional e por algoritmos que estão se tornando cada vez mais avançados, o que transforma a IA em uma ciência que interage com áreas como estatística, matemática, linguística, neurociência e até filosofia (LUDERMIR, 2021).

De acordo com Gabriel Filho (2023), as estratégias principais utilizadas em Inteligência Artificial são baseadas em modelos matemáticos e computacionais projetados para replicar o raciocínio e o aprendizado humano. Entre os modelos famosos, encontramos as redes neurais artificiais, que imitam de forma simplista o funcionamento dos neurônios no cérebro humano. As redes neurais profundas, que usam múltiplas camadas de processamento, têm a capacidade de identificar padrões complexos em imagens, sons e textos de maneira extremamente precisa. Um exemplo dessa tecnologia são os sistemas de reconhecimento facial, que examinam milhares de características em uma imagem para reconhecer uma pessoa com clareza.

Outra abordagem bastante comum para Bigonha (2025), é a árvore de decisão, a qual organiza dados em formato de gráficos de fluxo, tornando mais simples a análise das escolhas feitas pelo algoritmo. Essa abordagem se destaca em setores como finanças, saúde e direito, onde é essencial que a razão por trás de cada decisão seja compreensível. Os algoritmos de agrupamento, como k-means, junto com técnicas de redução de dimensões como PCA, são utilizados para identificar padrões ocultos em grandes conjuntos de dados, sendo aplicáveis em áreas como marketing, biologia e segurança da informação. Além disso, métodos de regressão logística, máquinas de vetores de suporte e modelos bayesianos também fazem parte do conjunto de ferramentas da Inteligência Artificial, cada um oferecendo benefícios em contextos particulares.

O aprendizado por reforço tem conquistado destaque recentemente, especialmente em contextos como jogos, como o AlphaGo, da DeepMind, e em sistemas de controle autônomo, como veículos e robôs



industriais. Nesse tipo de abordagem, o agente inteligente se envolve com o ambiente e aprende a tomar decisões baseado em feedback contínuo, tanto positivo quanto negativo, até que suas táticas sejam aperfeiçoadas. Esse modelo de aprendizado é promissor para situações que requerem uma tomada de decisão flexível e adaptável, sendo cada vez mais adotado em robótica, logística e sistemas de recomendações personalizadas (SCHUSSLER et al. 2018).

Além das técnicas em si, Gabriel Filho (2023) pontua que, é crucial ressaltar a importância das ferramentas e bibliotecas de programação, como TensorFlow, PyTorch, Scikit-learn e Keras, que tornaram a construção de modelos de Inteligência Artificial mais acessível. Essas plataformas disponibilizam módulos prontos para diversas tarefas, como classificação, regressão, processamento de linguagem natural e visão computacional, diminuindo a dificuldade técnica para a aplicação da IA em vários campos do conhecimento. Com esses recursos, instituições acadêmicas, startups e até agências governamentais têm conseguido criar soluções inteligentes que atendem suas demandas.

Embora frequentemente Inteligência Artificial e Aprendizado de Máquina sejam considerados equivalentes, Bazeia (2023) cita que, é importante compreender suas distinções e como se relacionam com outras tecnologias emergentes, como Big Data, Internet das Coisas, Computação em Nuvem e Blockchain. A IA depende fundamentalmente de grandes volumes de dados para aprimorar seus algoritmos, enquanto a IoT facilita a geração contínua dessas informações através de sensores em dispositivos móveis, veículos e equipamentos industriais. Por sua vez, a computação em nuvem oferece a infraestrutura necessária, de alto desempenho e escalável, para processar e armazenar os dados que são utilizados no treinamento dos modelos.

Para Arão (2024), outra tecnologia que está se integrando cada vez mais à inteligência artificial é o Blockchain, particularmente em relação à segurança e ao rastreamento das decisões feitas por algoritmos. A fusão de IA com Blockchain tem sido aplicada para desenvolver sistemas mais seguros e que possam ser auditados, sendo capazes de documentar decisões em tempo real e assegurar a integridade dos processos automáticos. Essa união se mostra especialmente valiosa em áreas como saúde e logística, onde a clareza e a permanência das informações são essenciais.

Além disso, Morais e Branco (2023) pontuam que, a IA e a Robótica se complementam, onde a inteligência artificial oferece o “cérebro” e os algoritmos para as tomadas de decisão, enquanto a robótica opera no espaço físico com sensores, atuadores e sistemas de controle. Exemplos desse tipo de integração incluem robôs cirúrgicos, drones para entrega e veículos que operam de forma autônoma. Frequentemente, a IA não apenas possibilita a realização de tarefas, mas também a adaptação delas ao contexto, através de aprendizado contínuo baseado nas condições ao redor.

Finalmente, compreender essas diferenças é fundamental para evitar mal-entendidos conceituais e para estimular uma incorporação crítica e bem-informada das tecnologias digitais. Segundo Arão (2024), a



IA não funciona sozinha, mas em colaboração com outras inovações tecnológicas, formando um ecossistema interconectado que transforma as estruturas econômicas, as relações de trabalho, os modos de consumo e até a maneira como as pessoas percebem a realidade. Ao entender os princípios da IA e do Aprendizado de Máquina, estabelece-se a base essencial para examinar suas aplicações com mais precisão, identificar suas limitações e refletir sobre as consequências de seu uso sem as devidas regulamentações.

2.2 APLICAÇÕES DA INTELIGÊNCIA ARTIFICIAL EM SETORES ESTRATÉGICOS

Segundo Bernasiuk e Sarlet (2024), o setor da saúde é uma das áreas que mais se beneficia com a utilização da Inteligência Artificial, principalmente no que se refere ao diagnóstico antecipado e à adaptação dos tratamentos. Ferramentas fundamentadas em IA têm se mostrado altamente eficazes na detecção de padrões em exames de imagem, como raios-X, tomografias e ressonâncias magnéticas, frequentemente com precisão igual ou até melhor do que a de médicos. Esses sistemas são alimentados com um grande volume de exames e conseguem notar alterações sutis que podem passar despercebidas por um profissional, ajudando de maneira crucial no diagnóstico de condições como câncer, pneumonia, doenças do coração e distúrbios neurológicos.

Além da análise de imagens, Amaro Jr et al. (2024) citam que, a IA também tem sido amplamente aplicada em previsões e na medicina sob medida. Ao combinar dados clínicos, genéticos e comportamentais, esses sistemas inteligentes são capazes de prever a probabilidade de um paciente desenvolver determinadas condições, indicar medicamentos mais adequados e sugerir tratamentos baseados em perfis de pacientes semelhantes. Isso representa um avanço considerável em direção à medicina personalizada, na qual o tratamento é ajustado de acordo com as características individuais de cada paciente, o que tende a elevar as taxas de sucesso e minimizar efeitos indesejados.

A administração hospitalar é outro setor seriamente influenciado pela IA. Para Leite et al. (2024), através de algoritmos de otimização e aprendizagem de máquina, é possível antecipar a necessidade de leitos, acompanhar o estoque de medicamentos, agendar cirurgias de forma mais eficaz e até prever surtos de doenças com base em informações epidemiológicas. Durante pandemias, como a de COVID-19, sistemas alimentados por IA foram empregados para monitorar a propagação do vírus, estimar a curva de contágio e informar ações de saúde pública, provando seu valor também na saúde coletiva.

Adicionalmente, chatbots e assistentes virtuais têm sido integrados em serviços de saúde para fornecer informações básicas, triagens iniciais e suporte psicológico. Embora esses sistemas ainda apresentem limites, eles marcam um progresso na ampliação do acesso à saúde, especialmente em áreas remotas ou onde há falta de profissionais. No entanto, esses avanços também trazem à tona questões éticas e legais, como a segurança dos dados dos pacientes e a responsabilidade em casos de erros de diagnóstico, exigindo uma regulação cuidadosa na utilização dessas tecnologias (AMARO JR et al. 2024).



Ainda, é fundamental ressaltar que o uso de inteligência artificial na área da saúde deve servir como um apoio à intervenção humana, e não como um substituto. Segundo Bernasiuk e Sarlet (2024), a tomada de decisões médicas abrange não apenas a análise técnica, mas também a empatia, o discernimento ético e a comunicação, aspectos que os sistemas ainda não conseguem reproduzir de forma profunda. Assim, o maior valor da inteligência artificial na saúde reside em expandir a eficácia dos profissionais, tornando-os mais informados e eficientes, sem remover o elemento humano que é crucial no atendimento.

No campo da segurança pública, Pinheiro (2024) cita que, a Inteligência Artificial tem sido empregada como um recurso para a prevenção e combate do crime. Em várias cidades ao redor do globo, sistemas de monitoramento inteligentes, fundamentados em visão computacional e reconhecimento facial, têm sido implementados para identificar suspeitos em tempo real, observar comportamentos anômalos em espaços públicos e supervisionar áreas de risco. Esses sistemas utilizam vastos bancos de dados e câmeras integradas, possibilitando a intersecção de dados e a emissão de alertas automáticos às autoridades competentes.

Uma outra área significativa da IA nesse campo de acordo com Nagata (2024), se revela nos sistemas de análise preditiva, que avaliam registros de crimes, informações socioeconômicas e padrões de comportamento para estimar a probabilidade de ocorrências criminais em regiões específicas. Essa ferramenta tem sido aplicada para guiar o patrulhamento policial, distribuir recursos de forma estratégica e até mesmo prever possíveis episódios de violência urbana. Apesar de seu potencial promissor, este uso levanta questões éticas importantes, especialmente sobre o viés algorítmico e o risco de perpetuar estereótipos sociais ou práticas de discriminação.

Além disso, a IA tem desempenhado um papel importante na área de investigações criminais. Programas de análise de big data e mineração de texto podem processar grandes quantidades de informações coletadas de celulares, redes sociais, câmeras de segurança e documentos oficiais para descobrir conexões entre suspeitos, mapear organizações criminosas e acelerar a resolução de casos difíceis. Ademais, algoritmos de processamento de linguagem natural podem auxiliar na interpretação de depoimentos e relatórios policiais, contribuindo para a identificação de inconsistências ou padrões importantes (PINHEIRO, 2024).

Nos presídios, Vieira e Barbuda (2024) pontuam que, a IA tem sido aplicada para observar o comportamento dos detentos, prevenir desavenças e aperfeiçoar a gestão de recursos. Sensores inteligentes e sistemas de análise de vídeo são capazes de identificar movimentos suspeitos, aglomerações ou cenários de risco, proporcionando uma resposta mais rápida das equipes de segurança. Contudo, assim como em outros setores, a utilização intensa de tecnologias de monitoramento pede um equilíbrio entre segurança pública e respeito aos direitos fundamentais, incluindo a privacidade e a presunção de inocência.

Diante dessas inovações, é essencial que as políticas públicas que envolvem segurança com base em



IA sejam formuladas com clareza, supervisão humana e controle social. Para Nagata (2024), a aplicação de tecnologias altamente invasivas sem a devida regulação democrática pode resultar em abusos de autoridade e acentuar desigualdades sociais. Portanto, é vital regulamentar essas atividades, assegurando que a tecnologia beneficie a cidadania, e não a repressão arbitrária ou preconceito institucionalizado.

De acordo com Oliveira et al. (2024), a indústria se destacou como uma das líderes no aproveitamento da Inteligência Artificial como um impulsionador da nova era conhecida como Indústria 4.0, a qual se sobressai pela combinação de sistemas ciberfísicos, internet das coisas e decisões autônomas. Neste novo cenário produtivo, a IA é aplicada para simplificar processos, antecipar avarias de máquinas, controlar estoques e adaptar a produção de maneira imediata, com base nas exigências e condições operacionais. Este novo modelo promove uma eficiência superior, economia de recursos e adaptações mais ágeis nas cadeias de produção.

Uma das mais notáveis utilizações da IA na indústria segundo Gonçalves et al. (2023), encontra-se na manutenção preditiva, que se utiliza de sensores e algoritmos para realizar um monitoramento constante do funcionamento das máquinas e prever falhas antes de acontecerem. Isso previne interrupções inesperadas, diminui gastos com manutenção corretiva e estende a durabilidade dos equipamentos. Os sistemas também possuem a capacidade de aprender com os dados de funcionamento passados, ajustando continuamente e de maneira autônoma seus modelos preditivos.

Outra área promissora para Estrela et al. (2024), é a utilização de robôs colaborativos, ou cobots, que trabalham em conjunto com os operadores humanos em atividades repetitivas, perigosas ou que requerem grande precisão. Dotados de sensores, visão computacional e algoritmos de IA, esses robôs são capazes de se adaptar ao ambiente e responder de maneira segura a comandos humanos. Isso não só aumenta a produtividade, mas também melhora as condições de trabalho, reduzindo acidentes e a fadiga dos operadores.

A logística também foi significativamente transformada pelo uso da IA, através de sistemas de roteamento inteligente, previsões de demanda e gestão automatizada de estoques. Grandes redes de varejo e empresas de transporte aplicam algoritmos para determinar rotas eficientes, antecipar períodos de alta demanda e ajustar automaticamente os centros de distribuição. Isso resulta em uma cadeia de suprimentos mais eficaz, econômica e menos vulnerável a erros humanos (GONÇALVES et al. 2023).

No entanto, essas inovações trazem consigo o desafio da substituição do trabalho tradicional, especialmente em funções operacionais e administrativas. Para Oliveira et al. (2024), a automação inteligente pode levar à perda de milhões de empregos, necessitando de uma requalificação imediata dos profissionais, além de políticas públicas que garantam uma transição para novas oportunidades de trabalho. Assim como em outros setores, o emprego responsável da IA na indústria requer um equilíbrio entre a inovação tecnológica e a inclusão social, assegurando que os benefícios da automação sejam amplamente



distribuídos.

2.3 DESAFIOS E RISCOS ASSOCIADOS AO USO DA INTELIGÊNCIA ARTIFICIAL

Apesar dos progressos notáveis, os sistemas de Inteligência Artificial ainda estão longe de serem infalíveis. De acordo com Bernasiuk e Sarlet (2024), um dos riscos técnicos mais significativos associados à utilização da IA é a possibilidade de erros ou decisões incorretas, que podem surgir por causa de falhas no código, treinamentos com dados tendenciosos ou limitações dos modelos em enfrentar situações não previstas. Tais falhas podem ter repercussões severas, sobretudo em áreas delicadas como saúde, aviação, justiça e transporte. Por exemplo, em diagnósticos médicos automatizados, um erro na classificação pode atrasar o tratamento de doenças graves; em carros autônomos, um erro na leitura dos sensores pode ocasionar acidentes letais. Desse modo, a confiabilidade da IA se baseia na solidez dos algoritmos e na qualidade dos dados utilizados.

Outro ponto crucial é o desafio da "caixa-preta", comum em sistemas de aprendizado profundo, que impede a compreensão das decisões dos algoritmos. Essa falta de clareza não apenas diminui a confiança dos usuários, mas também restringe a habilidade de auditar e corrigir erros quando eles surgem. Segundo Amaro Jr et al. (2024), em campos como direito e medicina, a justificativa das decisões é essencial. Assim, a falta de explicações gera preocupações sérias em termos éticos e legais em relação à delegação de decisões complexas para máquinas automatizadas. A necessidade de explicações algorítmicas e o surgimento de técnicas de IA interpretáveis (Explainable AI – XAI) aparecem como soluções para esse problema técnico.

Adicionalmente, Barroso (2024) cita que, muitos sistemas de IA são extremamente vulneráveis ao ambiente em que atuam. Mudanças sutis nos dados de entrada — como uma imagem levemente modificada ou informações incompletas — podem resultar em respostas inesperadas ou totalmente incorretas. Isso evidencia a fragilidade estrutural desses sistemas em contextos que exigem flexibilidade e bom senso, características ainda bastante humanas. Portanto, a redução dos riscos técnicos depende da contínua validação dos sistemas, da supervisão humana e do desenvolvimento de estratégias que permitam aos algoritmos lidar com incertezas de maneira mais eficaz.

A adoção massiva de soluções de inteligência artificial provoca mudanças significativas no mercado de trabalho, alimentando um dos principais receios da sociedade moderna: o desemprego gerado pela tecnologia. Segundo Carvalho (2021), trabalhos repetitivos e padronizados, principalmente nas áreas operacionais, administrativas e de atendimento ao cliente, estão sendo gradualmente substituídos por sistemas automatizados, que oferecem maior rapidez, precisão e economia. Isso afeta diretamente as oportunidades de emprego de milhões de trabalhadores, especialmente aqueles com baixa escolaridade ou acesso limitado a formação tecnológica. Como resultado, as desigualdades sociais e econômicas tendem a se intensificar, especialmente em nações em desenvolvimento.



Além da questão do desemprego, a disseminação da inteligência artificial evidencia o problema da exclusão digital de acordo com Peixoto e Paiva (2024). O acesso a ferramentas tecnológicas, treinamento em habilidades digitais e oportunidades para requalificação não ocorre de maneira justa entre diferentes regiões, classes sociais e idades. As comunidades mais vulneráveis, como habitantes de áreas rurais, idosos e pessoas em condição de pobreza, costumam ser excluídas do processo de inovação, o que cria obstáculos adicionais para inclusão e participação cidadã. A centralização do conhecimento técnico e do poder computacional nas mãos de um pequeno número de grupos econômicos e países desenvolvidos contribui para ampliar o fosso tecnológico global.

A abordagem a esses problemas exige a criação de políticas públicas que priorizem a educação digital, a inclusão tecnológica e a requalificação profissional. Segundo Barroso (2024), governos, empresas e instituições educacionais devem colaborar para assegurar que os benefícios da inteligência artificial sejam acessíveis a mais do que apenas uma elite técnica e econômica. A transição para uma economia cada vez mais automatizada deve ser acompanhada de investimentos em treinamento, redes de proteção social e mecanismos que redistribuam oportunidades. Se isso não ocorrer, a inteligência artificial, em vez de fomentar progresso coletivo, pode reforçar cenários de exclusão e desigualdade social ainda mais acentuados.

Para Aranha e Costa (2023), um dos pontos mais importantes e discutidos sobre a utilização da Inteligência Artificial são os preconceitos algorítmicos, ou seja, a tendência dos sistemas de manter ou até intensificar os preconceitos que já existem nos dados usados para seu treinamento. Esses preconceitos podem se manifestar de maneira sutil e não intencional, mas têm efeitos sérios e reais em decisões acerca de crédito, admissão em instituições de ensino, julgamentos, diagnósticos na área da saúde e seleção de funcionários. Quando os dados históricos contêm discriminações de raça, gênero, classe ou outras, os algoritmos aprendem e replicam esses padrões, reforçando desigualdades estruturais sob a fachada de uma neutralidade técnica.

Exemplos notáveis como o sistema COMPAS, conforme citado com Peixoto e Paiva (2024), utilizado na justiça dos Estados Unidos, mostraram como algoritmos podem classificar réus negros como mais propensos a reincidir em comparação com réus brancos que têm o mesmo histórico. Na área da saúde, pesquisas apontam que modelos de triagem médica costumam priorizar menos pacientes negros devido a associações enviesadas entre etnia e risco clínico. Essas distorções evidenciam que a IA não está livre dos problemas sociais presentes nos dados que a sustentam. Pior ainda: devido ao fato de serem decisões automatizadas, localizar e corrigir esses vieses se torna ainda mais complicado.

A redução dos preconceitos requer segundo Carvalho (2021), em primeiro lugar, uma análise crítica na elaboração e validação dos sistemas de IA, com a colaboração de equipes com diferentes especialidades que levem em conta aspectos técnicos, sociais e éticos. É essencial realizar auditorias regulares sobre os



dados e algoritmos e implementar métricas de equidade que considerem o impacto das decisões automatizadas em diversos grupos sociais. Além disso, a legislação deve acompanhar essas iniciativas, estabelecendo diretrizes claras para a responsabilização em situações de discriminação algorítmica. O enfrentamento dos preconceitos na IA não é apenas um desafio técnico, mas também um compromisso com a justiça social e os direitos humanos.

A crescente incorporação da Inteligência Artificial nos setores produtivos, administrativos e sociais suscita também inquietações relacionadas à dependência tecnológica e aos perigos à soberania digital. De acordo com Barroso (2024), nações que carecem de uma infraestrutura tecnológica própria ou da capacidade de desenvolver sistemas de IA tendem a se transformar em meros consumidores das soluções elaboradas por grandes nações tecnológicas, como os Estados Unidos e a China. Isso provoca uma desigualdade em termos de poder e conhecimento que prejudica a autonomia decisória de países periféricos e acentua sua fragilidade geopolítica.

A centralização do controle da IA em um pequeno número de grandes empresas, conhecidas como "Big Techs", como Google, Amazon, Microsoft e Alibaba, resulta em uma concentração que afeta não apenas a economia, mas também o campo político e informativo. De acordo com Rocha et al. (2024), esses grupos dominam grandes volumes de dados, possuem os algoritmos mais avançados e estabelecem normas tecnológicas que influenciam comportamentos sociais, escolhas de consumo e até mesmo políticas públicas. Quando estados, empresas e cidadãos passam a depender dessas estruturas privadas para funcionar, o risco de prejudicar a soberania e a segurança digital nacional aumenta.

Para superar essa dependência, é necessário realizar investimentos estratégicos em pesquisa, inovação e em infraestrutura tecnológica nacional conforme aponta Barroso (2024). Nações que buscam preservar sua autonomia na era digital precisam incentivar a formação de ecossistemas locais de IA, apoiar instituições acadêmicas e centros de pesquisa, promover startups tecnológicas e criar marcos regulatórios que favoreçam a soberania dos dados. A implementação de padrões abertos, a promoção da interoperabilidade e o fortalecimento da governança digital são ações essenciais para assegurar que a IA favoreça o fortalecimento, e não a fragilização, da soberania dos povos e das nações.

2.4 IMPLICAÇÕES ÉTICAS E REGULATÓRIAS

De acordo com Dourado e Aith (2022), a privacidade se destaca como uma das preocupações mais críticas no debate ético em torno da Inteligência Artificial, principalmente porque os algoritmos de aprendizado de máquina necessitam de grandes volumes de dados para funcionar de forma eficaz. Grande parte desses dados é pessoal e muitas vezes delicada, abrangendo informações como saúde, localização, hábitos de consumo, histórico de navegação e até características biométricas. O uso sem consentimento adequado configura uma transgressão direta ao direito à privacidade e à autodeterminação informativa,



garantidos por várias legislações em todo o mundo. No Brasil, a Lei Geral de Proteção de Dados (LGPD) representa um progresso significativo ao estabelecer diretrizes como finalidade, necessidade e segurança, limitando o tratamento de dados por entidades públicas e privadas. Contudo, a aplicação desta lei ao contexto da IA ainda necessita de interpretações e adaptações específicas que levem em consideração as complexidades da decisão automatizada.

Além da coleta e uso dos dados, Costa e Aranha (2023) citam que, a transferência e armazenamento de informações sensíveis também geram sérios dilemas éticos, especialmente quando esses dados são compartilhados com terceiros ou utilizados para propósitos diferentes daqueles informados inicialmente ao titular. Plataformas digitais que utilizam IA frequentemente reúnem e conectam dados de várias origens, o que pode resultar na formação de perfis comportamentais extremamente detalhados e invasivos, sem que o usuário tenha total consciência sobre o que está sendo feito com suas informações. Essa realidade requer não apenas o fortalecimento de mecanismos de consentimento claro e informado, mas também o desenvolvimento de auditorias independentes e sistemas de governança que garantam a rastreabilidade e controle dos dados pessoais utilizados em sistemas de IA.

A automação das decisões que antes eram realizadas somente por seres humanos traz à tona um desafio significativo no âmbito legal e ético: quem deve arcar com a responsabilidade em casos de erro, falha ou prejuízo causado por um sistema de inteligência artificial? Essa questão se torna ainda mais complicada quando se trata de um sistema que se adapta autonomamente, ou seja, que aprende ao longo do tempo e faz escolhas baseadas em dados que nem mesmo aqueles que o desenvolveram conseguem prever de maneira completa. De acordo com Rocha et al. (2024), a responsabilidade passa a ser difusa, abrangendo programadores, provedores de dados, empresas que utilizam a tecnologia e, em certas situações, até os próprios usuários finais. Essa complexidade indica a necessidade de revisitar conceitos tradicionais do direito civil, como culpa, dolo e nexos causal, para torná-los compatíveis com as novas realidades tecnológicas.

Neste cenário, Costa e Aranha (2023) pontuam que, alguns especialistas sugerem a criação de novas figuras jurídicas específicas, como a "personalidade digital", que poderia atribuir responsabilidade objetiva a certos tipos de sistemas autônomos. Outros argumentam que tanto os fabricantes quanto os operadores devem ser responsabilizados de forma solidária, semelhante à responsabilidade por produtos com defeito. O aspecto crucial é assegurar a presença de mecanismos claros e eficazes para reparação de danos, principalmente em setores delicados como saúde, finanças e justiça. A falta de uma estrutura regulatória específica relacionada à responsabilidade por decisões automatizadas pode resultar em uma lacuna perigosa, onde as vítimas de erros de inteligência artificial ficam desprovidas de assistência e os responsáveis evitem prestar contas.

A clareza nos sistemas de inteligência artificial é uma demanda ética essencial para que exista



confiança pública nas escolhas automáticas. Contudo, muitos algoritmos, em particular os que utilizam redes neurais profundas, funcionam como verdadeiras "caixas-pretas", gerando resultados sem que possamos entender os critérios lógicos por trás deles. Essa falta de transparência prejudica não só a supervisão institucional, mas também os direitos dos cidadãos afetados, que ficam incapazes de contestar ou entender as decisões que os envolvem. A capacidade de explicação, por conseguinte, se torna um aspecto vital da justiça algorítmica, permitindo que os indivíduos impactados por decisões da IA compreendam como e por que estas foram feitas (FORNASIER E KNEBEL, 2020).

Entretanto, ainda segundo Fornasier e Knebel (2020), criar modos de explicação não é uma tarefa simples. Existe uma tensão constante entre a eficiência técnica do modelo e sua facilidade de entendimento. Modelos que são mais precisos tendem a ser menos compreensíveis, enquanto os mais claros podem apresentar desempenho inferior. No entanto, é imprescindível buscar um equilíbrio que proporcione algum nível de auditoria e entendimento das decisões automatizadas, especialmente em áreas de grande impacto. Projetos de "IA explicável" (Explainable AI – XAI) já estão sendo desenvolvidos, mas ainda precisam de normas e regulamentações adequadas. Uma governança ética sólida deve assegurar que os sistemas de IA sejam, no mínimo, passíveis de auditoria por entidades reguladoras ou especialistas externos, a fim de garantir a conformidade com direitos fundamentais.

A falta de normas regulatórias específicas para a Inteligência Artificial é um dos maiores desafios para o desenvolvimento de uma IA que seja ética, segura e acessível a todos conforme citam Dourado e Aith (2022). Apesar de existir uma variedade de legislações, como o Código de Defesa do Consumidor, o Marco Civil da Internet e a LGPD, essas não consideram a complexidade da IA e, portanto, possuem lacunas significativas. A regulamentação da IA deve se estender além da proteção de dados e dos direitos do consumidor; deve incluir princípios como justiça algorítmica, transparência, supervisão humana e responsabilidade antecipada. A União Europeia já está se movendo nesse caminho com sua proposta de regulamentação da IA, que define categorias de risco e requisitos proporcionais, funcionando como um modelo para outras nações.

Em complemento, Costa e Aranha (2023) citam que, no Brasil, ainda há muito a ser realizado para criar um sistema normativo claro e abrangente sobre a questão. Existem propostas legislativas em tramitação no Congresso Nacional, mas carecem de uma interligação adequada e de um debate público mais amplo. A governança da IA deve ir além de apenas diretrizes legais: deve incorporar comitês interdisciplinares, envolvimento da sociedade, avaliações técnicas, princípios éticos e iniciativas educacionais. Apenas com uma estratégia colaborativa e integrada será possível assegurar que a inteligência artificial sirva como um meio para o avanço humano, ao invés de se tornar um foco de riscos incontroláveis. Assim, o desenvolvimento de um marco regulatório eficaz é uma necessidade urgente, tanto política quanto ética e técnica, da sociedade moderna.



3 CONCLUSÃO

A ascensão da Inteligência Artificial e do Aprendizado de Máquina representa uma transformação profunda na forma como a sociedade opera, afetando diretamente áreas cruciais como saúde, segurança e indústria. A habilidade de analisar grandes quantidades de dados rapidamente e identificar padrões úteis para a tomada de decisões tem permitido progresso que antes parecia impossível, aprimorando diagnósticos médicos, aumentando a vigilância em tempo real e automatizando processos industriais de maneira mais eficiente. No entanto, essas vantagens não vêm sem questões éticas, políticas e legais. O potencial da IA, apesar de promissor, também apresenta vários desafios que demandam análise cuidadosa e regulação atenta.

O dilema contemporâneo mais significativo é a necessidade de encontrar um equilíbrio entre inovação tecnológica e os direitos fundamentais dos indivíduos. Problemas como a falta de transparência nos algoritmos, a possibilidade de erros nas decisões, o aumento das desigualdades sociais e a invasão da privacidade são preocupações reais e em crescimento. A complexidade das tecnologias de IA frequentemente dificulta a responsabilização por possíveis prejuízos, gerando insegurança legal e tornando urgente a necessidade de elaborar normas específicas que atendam aos diferentes níveis de risco associados a cada uso. A gestão da IA deve, portanto, ser considerada não apenas como um aspecto técnico, mas como um processo que envolve um diálogo abrangente entre ciência, legislação, ética e a sociedade civil.

Outro aspecto crucial é a necessidade de adotar uma abordagem inclusiva e democrática na criação de políticas públicas relacionadas ao tema. O avanço da IA não pode ser monopolizado por grandes empresas ou nações tecnologicamente superiores, sob pena de acentuar desigualdades tanto globais quanto locais. É vital que haja investimento público em pesquisa, formação de profissionais, apoio a startups regionais e incentivo a soluções tecnológicas que atendam às necessidades sociais. Simultaneamente, é essencial assegurar que os benefícios da IA sejam distribuídos de forma justa, combatendo a exclusão digital e promovendo a capacitação dos cidadãos para interagir com as novas tecnologias.

Finalmente, o futuro da Inteligência Artificial está mais relacionado com a maneira como a sociedade decide incorporá-la ao dia a dia do que com sua habilidade técnica. Isso não requer apenas inovações, mas também um compromisso coletivo e adherence a princípios democráticos. Para desenvolver uma Inteligência Artificial ética e segura, é essencial promover a educação digital, garantir a clareza nos algoritmos, realizar uma supervisão rigorosa e, principalmente, estabelecer regulamentações que assegurem que a tecnologia sirva ao bem coletivo. Apenas com esses fundamentos será viável estabelecer uma coexistência saudável entre seres humanos e máquinas, convertendo o poder da inteligência artificial em uma ferramenta de justiça, inclusão e desenvolvimento sustentável.



REFERÊNCIAS

- AMARO JR, E.; NAKAYA, H.; RIZZO, L. V. Inteligência artificial em saúde. Revista USP, São Paulo, n. 141, p. 41-50, abril/maio/junho 2024.
- ARÃO, Cristian. Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina. Fluxo Contínuo, v. 47, n. 3, 2024.
- BARROSO, L. R. Inteligência artificial: promessas, riscos e regulação. Rev. Direito Práx. 15 (04), 2024.
- BAZEIA, G. P. Fundamentos da inteligência artificial: o caso das redes generativas (GANS): dilemas éticos. 2023. 23f. Artigo (Sistemas de informação) - UNISUL. Florianópolis, 2023.
- BERNASIUK, H. L. B.; SARLET, G. B. S. Inteligência artificial e saúde: os impactos da IA generativa na saúde do Brasil. R. Dir. Gar. Fund., Vitória, v. 25, n. 2 p. 51-84, maio/ago. 2024.
- BIGONHA, R. S. Fundamentos do aprendizado de máquina. 1ª ed. Belo Horizonte, MG, 2025.
- CARVALHO, A. C. P. L. F. Inteligência artificial: riscos, benefícios e uso responsável. Estudos Avançados, v. 35, n. 101, 2021.
- COSTA, F. D. C.; ARANHA, E. Questões atuais sobre inteligência artificial: Riscos e emergência regulatória. Revista Acadêmica, vol. 96, n. 2, 2023.
- DOURADO, D. A.; AITH, F. M. A. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. Rev Saude Publica. 2022;56:80.
- ESTRELA, H. D.; ARAUJO, F. R. C.; MOREIRA, K. O. et al. A influência da inteligência artificial (IA) nos processos de automação industrial. EditoraES, v. 12, n. 3, 2024.
- FORNASIER, M. O.; KNEBEL, N. Inteligência artificial: desafios e riscos ético-jurídicos. Direito & Paz, ano XIV, n. 43, 2020.
- GABRIEL FILHO, O. Inteligência artificial e aprendizagem de máquina: aspectos teóricos e aplicações. São Paulo: Blucher, 2023.
- GIL, Antonio Carlos. Métodos e técnicas de pesquisa social. 6. ed. São Paulo: Atlas, 2008.
- GONÇALVES, L. S.; FANTAZIA, G. S.; OLIVEIRA, D. S. et al. Inteligência artificial na indústria 4.0. e-Acadêmica, v. 4, n. 2, e2642485, 2023.
- LEITE, K. R. V.; PASQUALI, M. A. B.; SILVA, T. A. S. N. et al. Transformação da saúde: o impacto da Inteligência Artificial na medicina moderna. Contribuciones a las ciencias sociales, v. 17, n. 5, 2024.
- LUDERMIR, T. B. Inteligência artificial e aprendizado de máquina: estado atual e tendências. Estudos avançados, v. 35, n. 101, 2021.
- MORAIS, F. D. B.; BRANCO, V. R.C. A Inteligência Artificial: conceitos, aplicações e controvérsias. XX Simpósio Internacional de Ciências Integradas da UNART, Guarujá, 2023.



NAGATA, S. V. Utilização da inteligência artificial na segurança pública e sua contribuição na Polícia Militar. *Brazilian Journal of Development*, v. 10, n. 6, 2024.

OLIVEIRA, Viary Santana de; SANTOS, L. de Freitas; FERREIRA, M. Porto. Inteligência artificial na automação de processos industriais e seus impactos. *Revista de Economia Mackenzie*, v. 21, n. 1, p. 162–182, 2024.

PEIXOTO, F. G.; PAIVA, E. L. A. Desafios éticos do uso de inteligência artificial no ensino básico. *Revista Caderno Pedagógico*, v. 21, n. 13, 2024.

PINHEIRO, F. M. L. O papel da tecnologia e da inteligência artificial na segurança pública: desafios e convergências com os direitos fundamentais individuais e sociais. *Revista do Ministério Público do Estado do Rio de Janeiro* n° 94, out./dez. 2024.

ROCHA, Uelisson Borges. SANTOS, Wagna Piler Carvalho dos. NANO, Rita Maria Weste. Uso da inteligência artificial: avanços, riscos e desafios relacionados à propriedade intelectual. *Revista Científica Multidisciplinar Núcleo do Conhecimento*. Ed. 06, Vol. 03, pp. 137-149, 2024.

SCHUSSLER, P. J.; BASTIANI, E.; BUSSLER, N. R. C. Inteligência artificial e aprendizado de máquina: utilizando o entendimento da inteligência humana para reprodução na computação. *Salão do Conhecimento*, Unijuií, 2018.

VIEIRA, Paulo Henrique Brant; BARBUDA, Alex Soares. A aplicação da inteligência artificial na atividade de inteligência de segurança pública. *Revista Jurídica do Nordeste Mineiro*, v. 8, n. 1, 2024.