



COMO A IA PODE FALHAR NA PARTE ÉTICA E MORAL

Maria Eduarda Medeiro Porto- 824144948¹

Matheus Alves Santana- 824144952

2025

A inteligência artificial (IA) tem nos últimos tempos assumido um papel relevante nos mais diversos setores da nossa sociedade. Estamos num ponto sem retorno, e o nosso futuro passará naturalmente pela incorporação da inteligência artificial, na nossa vida diária, seja profissional, seja pessoal. A ideia da existência de máquinas “pensantes” e que tomem decisões pelos Humanos levanta uma série de questões éticas que devem estar presentes aquando do desenvolvimento e incorporação da inteligência artificial nos mais diversos setores da sociedade. É fundamental estudar e investigar as melhores abordagens à sua integração. A presente pesquisa identifica, os principais princípios orientadores da ética e moral, na área da informática e sistemas de informação, em contexto de uso de sistemas inteligentes e autônomos.

1 A ética e a Inteligência Artificial

A Inteligência Artificial (IA) tem se tornado uma parte integral da sociedade moderna, influenciando decisões em diversos setores, como saúde, finanças e transporte.

No entanto, a implementação de IA levanta preocupações éticas e morais significativas, que podem resultar em falhas e consequências indesejadas.

O estudo da Ética na Inteligência Artificial (IA) surgiu da necessidade de examinar os danos que as tecnologias de IA podem infligir tanto a indivíduos quanto à sociedade. Mesmo que esses danos ocorram de maneira involuntária, ou seja, sem a intenção deliberada de desenvolver aplicativos que comprometam a privacidade dos usuários ou que apresentem discriminação ou viés, essas soluções podem ter repercussões significativas em áreas como trabalho, educação e saúde.

¹ mariaeduarda.portom@gmail.com

A Ética na IA abrange um conjunto de princípios e valores que orientam o desenvolvimento, uso e implementação da inteligência artificial, enfatizando a importância da responsabilidade, justiça, segurança e transparência.

Um dos principais desafios consiste em assegurar que os algoritmos sejam justos e não perpetuem preconceitos ou discriminações.

Isso pode ser alcançado por meio de um processo de curadoria e organização dos dados, análise cuidadosa dos dados utilizados para o treinamento dos modelos de IA e a implementação de medidas para mitigar possíveis vieses. A seguir, são descritos casos em que se tornou explícita a necessidade de tratarmos a ética na IA e o uso da IA ética de forma sistematizada.

Vieses de gênero, raça e nacionalidade em tecnologias de tradução automática e processamento de linguagem natural. Demonstraram que a ferramenta de tradução automática mais utilizada no mundo, Google Translate, pode exibir vieses de gênero e uma forte tendência para defaults/padrões masculinos, embora os autores tenham demonstrado que esses vieses resultam dos dados do mundo real utilizados no treinamento do sistema de tradução automática e há uma provável relação com a forma como a sociedade fala e escreve sobre gênero no mundo do trabalho.

-Vieses na seleção de pessoas por departamentos de recursos humanos que utilizam as hiring platforms ou plataformas tecnológicas de contratação de pessoal.

-Vieses raciais em algoritmos de IA que classificaram pessoas como não sendo seres humanos). Conforme relatado em 2015, Jacky Alcine, enquanto utilizava o sistema conhecido como Google Photos, percebeu que o aplicativo de reconhecimento de rostos (faces) rotulava ele e seu amigo como gorilas.

2 Solução dos desafios éticos e morais envolvendo as IA's

Para responder a esses desafios, analisei a proposta de pesquisadores que propuseram o uso de abordagens que combinem e integrem as abordagens de IA baseadas em redes neurais (que são eficientes para aprendizado sobre um grande volume de dados) com as abordagens da IA que utilizam a lógica simbólica (que permite a formulação rigorosa do processo de raciocínio). Esses sistemas, denominados inteligência artificial neurosimbólica, permitem, em princípio, interpretar e explicar de forma lógica e rigorosa o comportamento dos sistemas de IA construídos sob a abordagem das redes neurais artificiais. Isso decorre do fato de sistemas lógicos, por construção, terem semântica rigorosamente definida, além de estabelecerem os fundamentos também rigorosos da inferência e do raciocínio e por permitirem a representação formal de normas e valores, que podem ser representados nos sistemas de IA, idealmente mais éticos.

Além disso, são necessários:

- **Desenvolvimento de Algoritmos Éticos:** É fundamental que os desenvolvedores de IA integrem princípios éticos desde a fase de design. Isso inclui a consideração de diversidade e inclusão nos dados de treinamento.
- **Transparência e Explicabilidade:** Criar sistemas que permitam a auditoria e a compreensão das decisões da IA. Isso pode incluir a implementação de técnicas de “IA explicável” que ajudem os usuários a entender como as decisões são tomadas.
- **Regulamentação e Normas:** Estabelecer diretrizes e regulamentações claras que orientem o uso ético da IA. Organizações como a ACM e a IEEE já estão desenvolvendo códigos de ética que podem servir como base.
- **Educação e Conscientização:** Promover a educação sobre ética em IA entre desenvolvedores, usuários e o público em geral. Isso pode ajudar a criar uma cultura de responsabilidade em torno do uso da tecnologia.