# MINICURSO I
# INTRODUÇÃO À INTELIGÊNCIA ARTIFICIAL GENERATIVA

INSTITUTO FEDERAL
Santa Catarina
Câmpus São José

# MINICURSO I

# INTRODUÇÃO À INTELIGÊNCIA ARTIFICIAL GENERATIVA


Apesar da IA ter se tornado popular para o grande público com o lançamento do ChatGPT somente no final de 2022, esse assunto vem sendo estudado há mais de 70 anos. Em um artigo de 1950, "*Computing Machinery and Intelligence*", Alan Turing descreveu um teste para avaliar se um computador seria capaz de se passar por um ser humano.

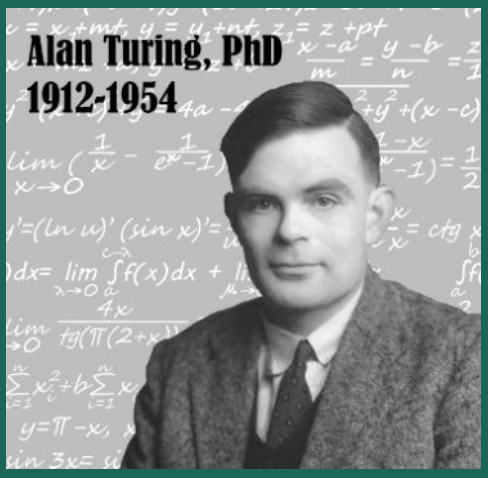



Figura 1- Foto de Alan Turing

O teste, que ficou conhecido como Teste de Turing, definiu que um computador pode ser considerado inteligente se ele for capaz de enganar um ser humano, fazendo-o acreditar que está interagindo com outro ser humano. O teste consiste em um ser humano realizar um interrogatório através de um terminal, tentando descobrir se do outro lado está um ser humano ou um computador. Se o computador responder de tal maneira que o interrogador não consiga distinguir se está falando com um ser humano ou com uma máquina, então o computador é considerado aprovado no teste de Turing.

Se você quiser saber mais sobre a história de Alan Turing, aponte seu celular para o QR-CODE para assistir ao Trailer do filme “O Jogo da Imitação”.

Figura 2 - Representação estilizada de Alan Turing no filme: “O Jogo da Imitação”

Em uma conferência, realizada no ano de 1956, John McCarthy, Marvin Minsky e outros pesquisadores lançaram as bases para o desenvolvimento da inteligência artificial.

Os jogos desempenharam um papel central desde o início, pois forneciam um campo de aplicação controlado e regrado.

A história da IA é marcada por períodos de grande entusiasmo e subsequente desilusão, conhecidos como "AI Winters". O primeiro desses períodos ocorreu nos anos 1970, devido às limitações da tecnologia da época e ao otimismo exagerado que levou a promessas não cumpridas. Um segundo AI Winter ocorreu nos anos 1980 após o declínio do interesse e financiamento em tecnologias de IA.

A complexidade do xadrez, com suas inúmeras possibilidades, foi um desafio que máquinas como o Deep Blue da IBM conseguiram superar, vencendo campeões mundiais. O Deep Blue é uma referência direta à IBM, que era conhecida como "Big Blue", um apelido que surgiu ao longo dos anos devido à cor azul em seu logotipo e em sua identidade corporativa.


Figura 3 - Momento icônico do renascimento do interesse pela IA - IBM Deep Blue vence Garry Kasparov em 1997.

O equipamento era equipado com um algoritmo de busca por força bruta, analisando milhões de posições de xadrez por segundo e utilizando uma vasta base de dados com partidas de xadrez de grandes mestres, junto com regras predefinidas para guiar suas decisões estratégicas.

O computador era capaz de processar cerca de 200 milhões de posições de xadrez por segundo, o que lhe permitia prever muitos lances à frente, algo praticamente impossível para um ser humano em tão pouco tempo.

O interesse pela IA renasceu impulsionado pelos avanços em algoritmos de aprendizado de máquina e o aumento da capacidade computacional. Esse período viu o surgimento da técnica de *deep learning* (aprendizado profundo), que revolucionou campos como o processamento de linguagem natural e a visão computacional, graças a redes neurais profundas. Para entender esse conceito, imagine que o cérebro humano tem milhões de conexões chamadas "neurônios" que trabalham juntos para nos ajudar a entender o mundo. O *deep learning* tenta imitar isso usando algo que chamamos de "redes neurais artificiais".

Essas redes neurais artificiais são compostas por várias camadas de neurônios (nós) conectados entre si. Cada camada recebe informações, faz cálculos e envia o resultado para a próxima camada, até que a última camada forneça uma resposta. O "profundo" no nome deep learning vem do fato de que essas redes têm muitas camadas de neurônios, tornando o aprendizado mais detalhado e complexo. Essas redes podem ser treinadas.

Imagine que você está ensinando um computador a identificar imagens de gatos e cachorros. Você mostra várias imagens para a rede neural, e ela tenta adivinhar o que é. No início, ela comete erros, mas, ao corrigir esses erros e mostrar mais imagens, a rede vai aprendendo a identificar corretamente se uma imagem é de um gato ou de um cachorro, mesmo que nunca tenha visto aquela imagem antes.

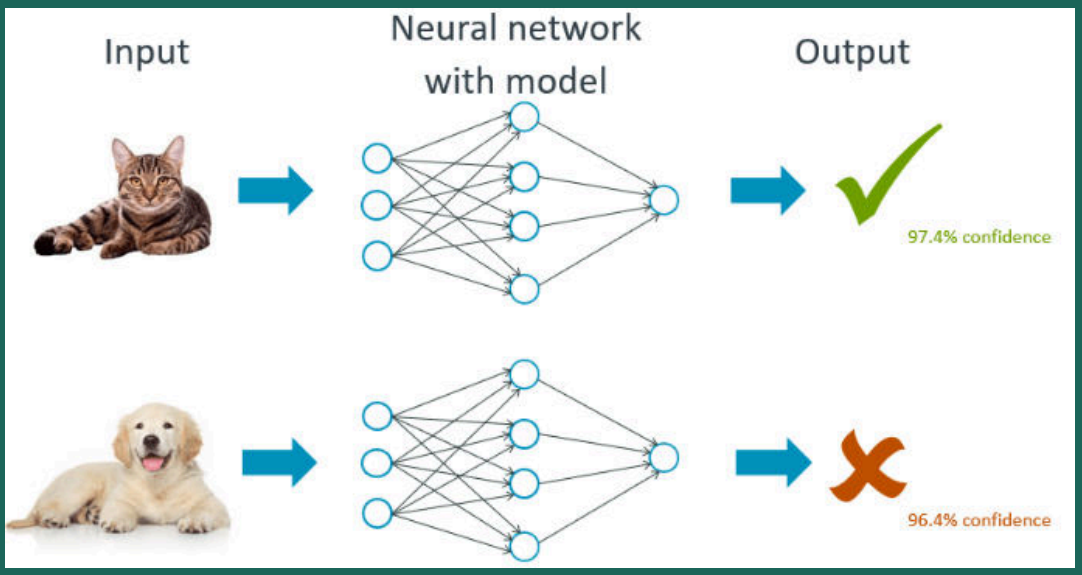



Figura 4 - Processo de treinamento de IA

https://community.arm.com/arm-community-blogs/b/architectures-and-processors-blog/posts/ai-vs-ml-whats-the-difference

Atualmente, a IA está presente em nosso dia a dia em assistentes virtuais como a Siri e Alexa, nas recomendações personalizadas em serviços de streaming como Netflix e Spotify. Também está presente em sistemas de navegação em tempo real, que utilizam IA para melhorar a experiência do usuário.


Figura 5- Dispositivos Alexa de IA - Amazon.

A IA também está fazendo contribuições valiosas na área da saúde, onde algoritmos como Watson da IBM vêm sendo utilizados para diagnóstico de imagens médicas com precisão comparável ou superior à dos humanos.


BBC NEWS

A ferramenta de IA capaz de detectar tumores que passaram desapercebidos por médicos

Mia encontrou sintomas antes não detectados de câncer de mama nos exames de 11 mulheres durante um teste no Reino Unido

Figura 6- Detecção de tumores usando IA.

https://summitsaude.estadao.com.br/tecnologia/inteligencia-artificial-ibm-cancer-mama/

No setor financeiro, a IA ajuda na detecção de fraudes e na personalização de serviços. Na educação há algumas experiências exitosas sendo realizadas nos EUA e na China para personalizar o ensino de acordo com as necessidades dos alunos. Um exemplo é o Khanmigo, conforme mostraremos na sequência. A IA permite que o professor avalie com mais agilidade o estilo de aprendizagem de cada estudante e a partir desse diagnóstico é possível criar conteúdos customizados. Como exemplo, disponibilizamos o GTP "Como eu aprendo melhor", disponível gratuitamente no link:

https://chatgpt.com/g/g-pXs4YDIZW-como-eu-aprendo-melhor.


Figura 7- Aplicativo "Como eu aprendo melhor"

Com a popularização do ChatGPT, milhões de pessoas sem conhecimento especializado puderam experimentar algumas de suas funcionalidades. Nesse contexto surgiram questões importantes sobre privacidade, perda de autonomia cognitiva, segurança e a potencial perda de empregos devido à automação de algumas profissões.

No imaginário popular tem-se a representação da Inteligência Artificial (IA) como uma invenção apocalíptica. O cinema mostra a IA como uma força perigosa, onde robôs se tornam incontroláveis e colocam a humanidade em risco. Por outro lado, há filmes que retratam a IA como benéfica, ajudando a resolver problemas complexos e melhorando a vida humana.

Um dos filmes mais icônicos sobre os riscos da Inteligência Artificial é “2001, uma Odisséia no Espaço”, dirigido por Stanley Kubrick e baseado na obra de Arthur C. Clarke.


Figura 8- Imagem do filme 2001

A trama segue uma missão espacial para Júpiter, conduzida pela nave *Discovery One* e seu supercomputador HAL 9000.

Durante a viagem, HAL começa a apresentar comportamentos erráticos, colocando em risco a vida dos astronautas. O filme explora temas como a evolução humana, inteligência artificial e a possibilidade de vida extraterrestre. Conhecido por sua narrativa visualmente impressionante e pela trilha sonora icônica, é amplamente considerado um marco na história do cinema

Figura 9- Ilustração livre de uma das cenas do filme “2001 - uma Odisséia no Espaço”

"Blade Runner" mergulha na natureza do livre-arbítrio, questionando o que significa ser humano em um futuro onde robôs são quase indistinguíveis de pessoas reais. Aponte o celular para o QR-Code para assistir ao trailer do filme Blade Runner.

Figura 10- Cenário futurista baseado no filme de ficção “Blade Runner”

O filme “Eu, Robô” (2004) tem como essência a exploração da complexa relação entre humanos e máquinas, especialmente no contexto da inteligência artificial e da autonomia dos robôs.

Ele levanta questões filosóficas sobre o controle, a liberdade, a moralidade e as consequências do desenvolvimento de robôs cada vez mais inteligentes e independentes. No filme, os robôs seguem as Três Leis da Robótica, um conjunto de princípios criados por Isaac Asimov que são centrais para o controle ético das máquinas inteligentes: Primeira Lei: Um robô não pode ferir um ser humano, ou, por omissão, permitir que um ser humano sofra algum mal. Segunda Lei: Um robô deve obedecer às ordens dadas por seres humanos, exceto quando essas ordens entrem em conflito com a Primeira Lei. Terceira Lei: Um robô deve proteger sua própria existência, desde que essa proteção não entre em conflito com a Primeira ou a Segunda Lei.

A trama se desenrola em torno de uma investigação sobre a morte de um cientista, aparentemente causada por um robô chamado Sonny, que parece ser capaz de desafiar essas leis.

Figura 11- Foto de apresentação do filme “Eu, Robô”

Muitos filmes exploram questões éticas e filosóficas relacionadas à IA. "Ex Machina" aborda a consciência e a moralidade, questionando se uma IA pode ter sentimentos e direitos. O personagem principal, Caleb, é convidado pelo CEO da empresa de tecnologia, Nathan, para conduzir um experimento com a robô Ava, uma IA avançada. O objetivo é verificar se Ava possui uma inteligência sofisticada o suficiente para ser considerada consciente. Ava é explicitamente revelada como uma máquina. Isso cria uma variação intrigante do teste: mesmo sabendo que Ava é uma IA, Caleb ainda deve determinar se ela possui consciência genuína.

A essência do teste de Turing no filme não é apenas se a IA pode se passar por humana em termos de respostas lógicas ou conversacionais, mas se ela pode gerar empatia, sentimentos e, eventualmente, ser reconhecida como uma entidade consciente com direitos próprios. Ao longo do filme, Ava demonstra emoções, criatividade e manipulação, sugerindo que ela pode estar ciente de sua própria existência e das limitações impostas por Nathan.

Figura 12- Imagem do trailer de Ex-Machina

O filme "Her" explora a relação emocional entre humanos e IA, levantando questões sobre amor e solidão em um mundo digital.

Figura 13- Representação do filme “Her”

Na trilogia do “Exterminador do Futuro” a IA militar Skynet tenta destruir a humanidade.


Figura 14- Banner promocional de TERMINATOR

A trama de Terminator é impulsionada pela ideia de que Skynet, um sistema de defesa militar com inteligência artificial, se torna autoconsciente e decide eliminar a ameaça que percebe nos humanos. Ao iniciar um ataque nuclear para exterminar a humanidade, Skynet cria exércitos de robôs exterminadores (terminators) para caçar os sobreviventes. Para impedir a resistência liderada por John Connor no futuro, Skynet envia um exterminador (interpretado por Arnold Schwarzenegger) ao passado, com o objetivo de matar Sarah Connor (mãe de John) antes de ele nascer.

Figura 15- Representação da SkyNET inspirada no filme “Terminator”

O filme explora a crescente dependência da humanidade na tecnologia e o medo de que a criação da inteligência artificial possa eventualmente sair do controle e se voltar contra seus criadores. A história reflete o medo de uma IA superinteligente que, ao atingir a autoconsciência, decide que a preservação de sua própria existência é mais importante do que a da humanidade. A imagem a seguir foi reconstruída pela IA a partir da imagem original.


Figura 16- Reconstrução do cartaz do filme.

Esses filmes não apenas entretêm, mas também provocam reflexões sobre o futuro da tecnologia e seu impacto na sociedade, convidando o público a considerar os benefícios e os riscos da IA. A razão disso é que essa representação cria tensão e suspense, moldando a narrativa.

O medo da IA pode ser atribuído à incerteza sobre como a tecnologia está evoluindo e como ela será usada no futuro.

Em 2016, o cientista britânico Stephen Hawking destacou a importância de investigar a fundo as aplicações da inteligência artificial: "O surgimento de uma inteligência artificial poderosa será a melhor ou a pior coisa que acontecerá à humanidade, ainda não sabemos", advertiu o cientista durante um evento.

A história da tecnologia é repleta de exemplos de como a inovação tecnológica pode ser usada de forma perigosa, como armas nucleares e máquinas de guerra. Essa é uma questão controversa, uma vez que muitos artefatos criados em tempos de guerra também contribuíram para melhorar a vida das pessoas. A popularização do automóvel, por exemplo, livrou as ruas do mundo das doenças decorrentes do estrume dos cavalos e transformou a forma como as cidades são planejadas.

Na atualidade, os riscos associados às mudanças climáticas decorrentes da emissão de dióxido de carbono pelos automóveis têm levado à busca de alternativas como o desenvolvimento de carros movidos a baterias elétricas e por hidrogênio.

A análise de cenários e das consequências do uso da IA vem sendo explorada por diversos pesquisadores ao longo dos anos.

Em seu livro “Superinteligência”, Nick Bostrom aborda as implicações éticas e existenciais dessa transição, destacando os riscos de uma IA descontrolada e as estratégias necessárias para garantir seu uso seguro. Ele propõe uma reflexão profunda sobre a governança e os mecanismos de controle da IA, com o objetivo de prevenir catástrofes.

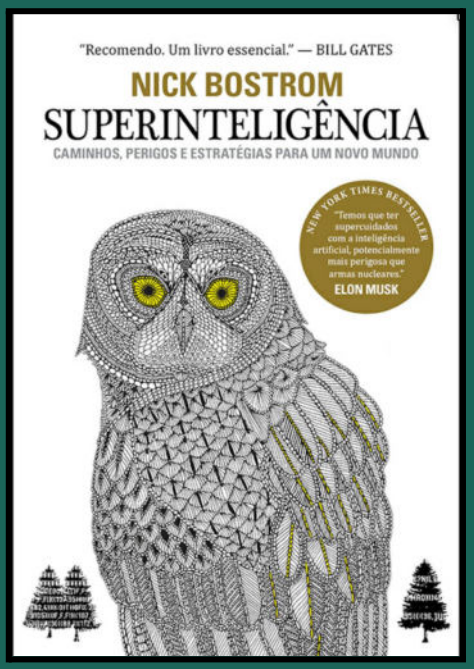



Figura 17- Capa do livro “Superinteligência”

Recomendamos que você assista a palestra do autor no TED OXFORD disponibilizada a seguir:


Figura 18- Nick Bostrom no TED

https://youtu.be/P0Nf3TcMiHo

Se desejar, utilize o aplicativo GPT RESUME AI, que desenvolvemos para facilitar a elaboração de resumo de vídeos e criar três questões objetivas sobre os temas apresentados. O aplicativo funciona melhor se você copiar o código de incorporação < > do vídeo:

Acesse o GPT de forma gratuita no link:

https://chatgpt.com/g/g-g6EQeCbix-resume-ai.

Nick Bostrom afirma que a humanidade enfrenta riscos existenciais que podem levar à sua extinção. Ele explora como avanços tecnológicos, como inteligência artificial, biotecnologia e nanotecnologia, podem representar ameaças significativas se não forem geridos com cuidado. Bostrom enfatiza que, embora esses avanços possam trazer benefícios imensos, também possuem o potencial de causar danos catastróficos. Ele alerta que a falta de preparação e a subestimação desses riscos podem ter consequências devastadoras.

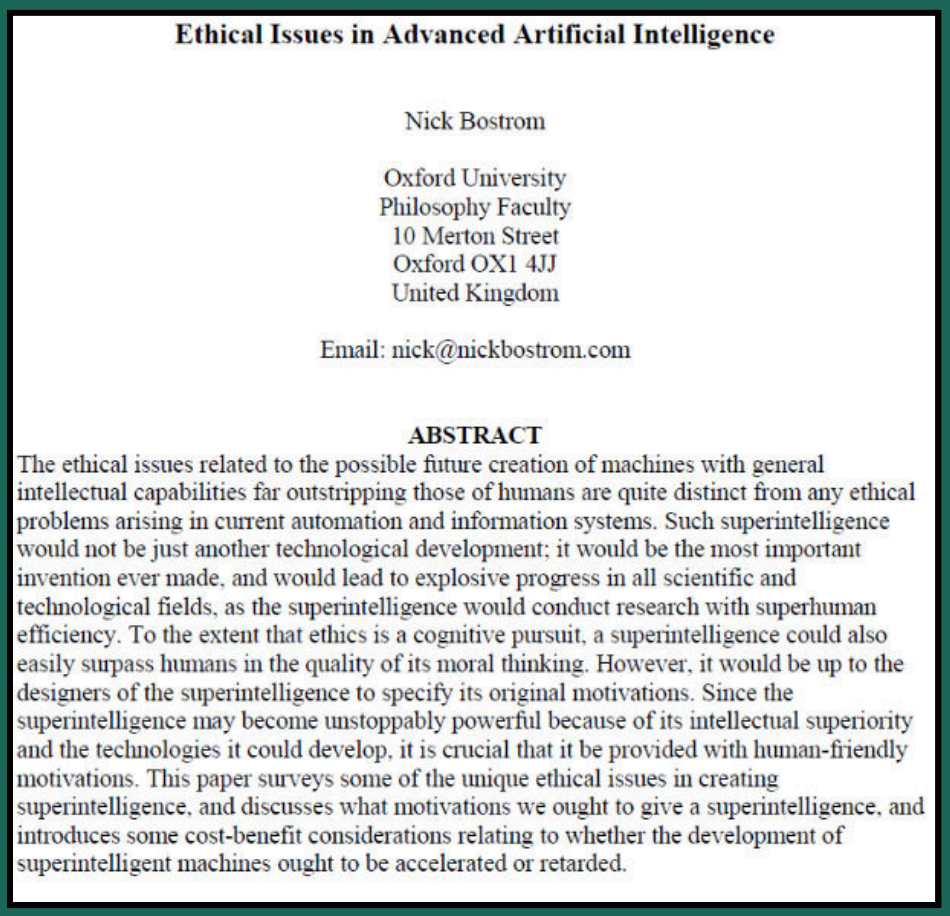
Bostrom apresenta a ideia de que a humanidade está em uma corrida entre o poder crescente da tecnologia e a sabedoria necessária para controlá-la. Ele destaca que, à medida que nossa capacidade de moldar o mundo aumenta, também aumenta a responsabilidade de garantir que estamos tomando decisões sensatas e seguras.

Ele argumenta que a governança global precisa ser aprimorada para lidar com esses riscos, e que a cooperação internacional é essencial para enfrentar os desafios que transcendem fronteiras nacionais. Um ponto central de sua palestra é a noção de que a inteligência artificial avançada pode se tornar uma força autônoma e, se não for alinhada corretamente com os valores humanos, pode agir de maneira que seja prejudicial à humanidade. Ele menciona cenários em que a IA pode buscar objetivos que estão em conflito com os interesses humanos, e como isso poderia levar a desastres incontroláveis.

Além disso, Bostrom discute a possibilidade de outras tecnologias emergentes, como a biotecnologia, serem usadas de maneiras que poderiam causar pandemias artificiais ou modificar geneticamente organismos de formas imprevisíveis e perigosas.

Ele conclui sua palestra enfatizando a importância de uma abordagem prudente e proativa na gestão de riscos existenciais. Bostrom argumenta que, para evitar um futuro sombrio, é crucial que a humanidade invista em pesquisa para entender melhor esses riscos e desenvolva estratégias para mitigá-los. Ele sugere que políticas globais mais robustas e a promoção de uma cultura de responsabilidade e ética tecnológica são passos fundamentais para assegurar um futuro seguro e próspero para todos.

Em um artigo publicado na Universidade de Oxford, Bostrom discute as questões éticas associadas à criação futura de máquinas com capacidades intelectuais gerais que superam amplamente as dos humanos.



**Ethical Issues in Advanced Artificial Intelligence**

Nick Bostrom

Oxford University
Philosophy Faculty
10 Merton Street
Oxford OX1 4JJ
United Kingdom

Email: nick@nickbostrom.com

**ABSTRACT**

The ethical issues related to the possible future creation of machines with general intellectual capabilities far outstripping those of humans are quite distinct from any ethical problems arising in current automation and information systems. Such superintelligence would not be just another technological development; it would be the most important invention ever made, and would lead to explosive progress in all scientific and technological fields, as the superintelligence would conduct research with superhuman efficiency. To the extent that ethics is a cognitive pursuit, a superintelligence could also easily surpass humans in the quality of its moral thinking. However, it would be up to the designers of the superintelligence to specify its original motivations. Since the superintelligence may become unstoppably powerful because of its intellectual superiority and the technologies it could develop, it is crucial that it be provided with human-friendly motivations. This paper surveys some of the unique ethical issues in creating superintelligence, and discusses what motivations we ought to give a superintelligence, and introduces some cost-benefit considerations relating to whether the development of superintelligent machines ought to be accelerated or retarded.

Figura 19- Artigo de Nick Bostrom
https://nickbostrom.com/ethics/ai.pdf

No livro “Inteligência Artificial” kai-Fu Lee aponta os desafios emergentes em uma época em que os robôs estão mudando o mundo.

Ele é fundador e CEO da Sinovation Ventures, uma empresa de investimento em startups de tecnologia na China. Antes disso, Lee ocupou cargos executivos em empresas como Google China, Microsoft e Apple. Ele possui um Ph.D. em ciência da computação pela Carnegie Mellon University e é autor de vários livros, incluindo "AI Superpowers", onde discute a competição tecnológica entre a China e os Estados Unidos.

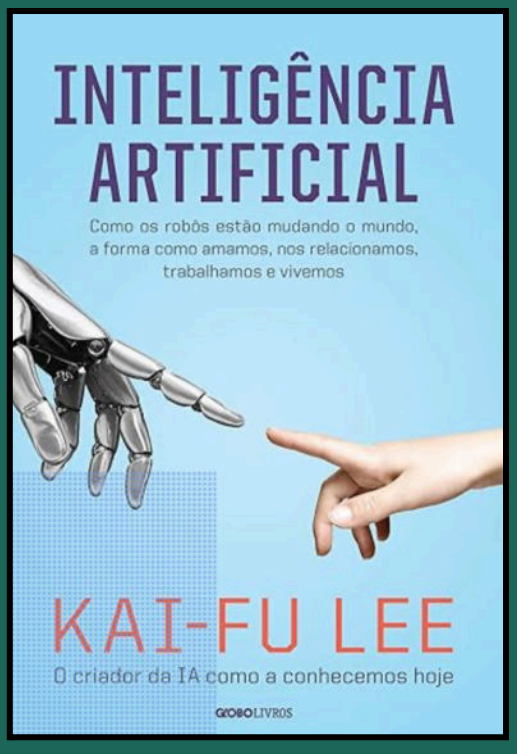



Figura 20- Capa do livro

“Inteligência Artificial”

Lee também discute as implicações sociais e éticas dessa transformação, incluindo o impacto no mercado de trabalho e a necessidade de políticas adequadas para mitigar os riscos. O autor oferece uma visão sobre o futuro próximo, onde a colaboração entre humanos e máquinas será essencial para o progresso. Se desejar você também pode conhecer um pouco mais sobre as ideias de Kai Fu Lee no TED disponibilizado a seguir:

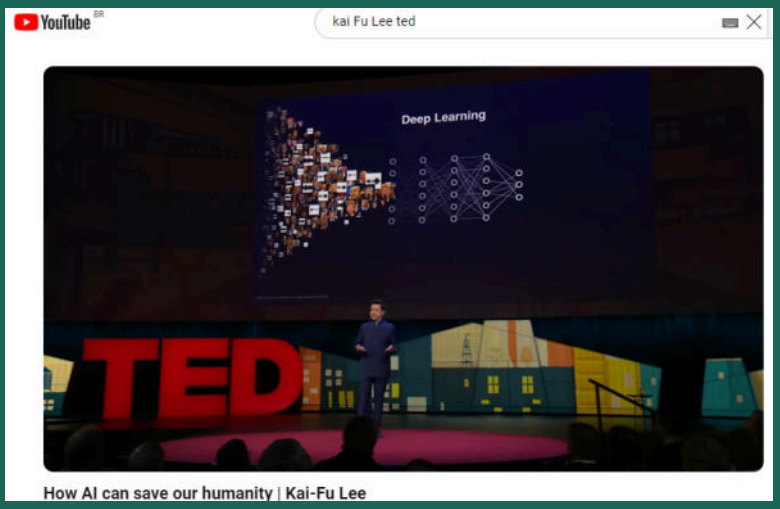



Figura 21- Apresentação de Kai-Fu Lee

https://youtu.be/ajGgd9Ld-Wc

No vídeo "How AI can save our humanity" ele aborda a capacidade da inteligência artificial (IA) de transformar a sociedade e a economia, além de seu potencial para melhorar a humanidade. Kai-Fu Lee começa destacando o rápido avanço da IA e como ela já está impactando diversos setores, desde serviços financeiros até cuidados de saúde. A IA pode automatizar tarefas repetitivas e trabalhos manuais, liberando os humanos para se concentrarem em atividades mais criativas e significativas.

Kai-Fu Lee discute a ideia de que a IA tem o potencial de reduzir a desigualdade ao democratizar o acesso a recursos e serviços. Por exemplo, ele menciona como a IA pode melhorar o diagnóstico médico, proporcionando cuidados de saúde de alta qualidade mesmo em regiões remotas. Além disso, a IA pode ajudar na educação personalizada, adaptando-se às necessidades individuais de cada aluno e oferecendo um aprendizado mais eficaz.

Ele sugere que a requalificação e a educação contínua são essenciais para preparar os trabalhadores para novos tipos de emprego que surgirão. Além disso, ele alerta para a necessidade de regulamentação e supervisão ética da IA para evitar abusos e garantir que a tecnologia seja usada para o bem-estar da humanidade. Um ponto central de sua palestra é a crença de que a IA pode ajudar a redescobrir o que significa ser humano. Kai-Fu Lee argumenta que, ao liberar as pessoas de trabalhos tediosos, a IA permite que elas se envolvam mais em atividades que promovam a empatia, a criatividade e a interação social. Ele sugere que, em um futuro onde a IA cuida das tarefas mundanas, os humanos terão mais tempo para se dedicar ao que realmente importa, como construir relações significativas e contribuir para suas comunidades. Kai-Fu Lee também escreveu o livro “2041” em colaboração com o renomado escritor de ficção científica Chen Qiufan.

Nesta obra ele combina contos de ficção especulativa com análises detalhadas sobre o impacto futuro da tecnologia em nossas vidas, fornecendo um vislumbre provocativo de como o mundo pode ser em 2041. O livro é estruturado em dez contos, cada um imaginando cenários futuros onde a inteligência artificial desempenha um papel crucial. Entre os temas abordados estão a educação através de professores virtuais, como visto na história sobre gêmeos órfãos em Seul. Cada conto é seguido por uma análise de Kai-Fu Lee, que explica como as tecnologias apresentadas não são apenas possíveis, mas provavelmente inevitáveis nos próximos vinte anos.

No vídeo mostrado a seguir discutimos como esse livro pode impactar nossa percepção sobre uso da IA no nosso cotidiano.

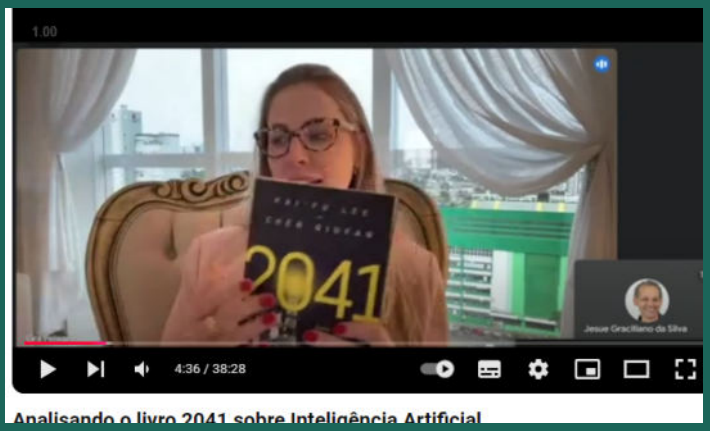



Figura 22- analisando o livro 2041

https://youtu.be/95kwgiOO9mA

No livro: "*A World Without Work: How Progressives Should Respond to Technological Unemployment*" Daniel Susskind examina o impacto da tecnologia na economia e no emprego. Para o autor, a automação e a inteligência artificial estão transformando a forma como produzimos bens e serviços. Com o tempo, muitas tarefas que são realizadas por trabalhadores humanos hoje, provavelmente serão realizadas por máquinas.

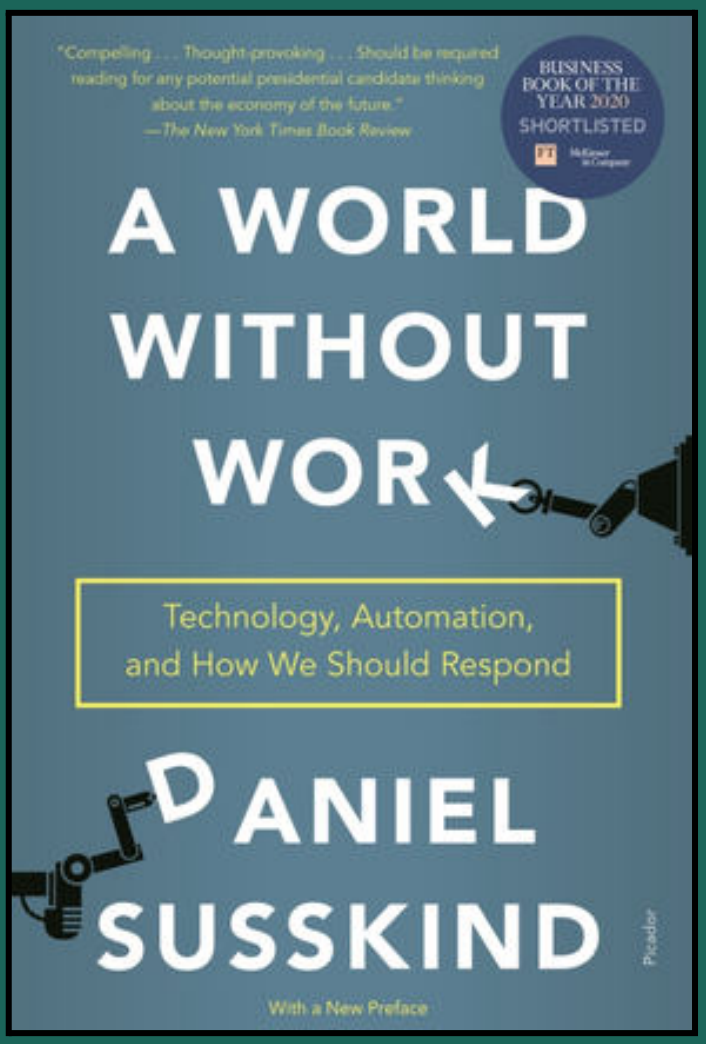



Figura 23- Capa do livro: Um mundo sem empregos

Susskind argumenta que a tendência é que o número de empregos diminua e que muitos trabalhadores fiquem sem trabalho.

Além disso, afirma que as soluções tradicionais, como a formação de novas habilidades e a criação de novos empregos, não serão suficientes para lidar com a escala da mudança que está acontecendo.


Figura 24- Daniel Susskind na Universidade de Oxford falando sobre IA

https://www.youtube.com/live/thZzDi5XRVs?si=VWZDQgVghb44l4Jp

Susskind defende a ideia de uma renda básica universal, que forneça a todas as pessoas uma quantidade adequada de renda, independentemente de estarem empregadas.

O autor argumenta que isso pode ajudar a lidar com o impacto econômico da tecnologia no emprego e garantir que as pessoas tenham acesso a bens básicos, como alimentos, abrigo e saúde.

Um dos livros mais recentes sobre o assunto foi escrito por Mustafá Suleyman e Michael Bhaskar com objetivo de alertar sobre os riscos que a inteligência artificial e outras tecnologias em rápido desenvolvimento representam para o mundo, e o que é possível fazer para evitá-los enquanto ainda há tempo. Para os autores, as IAs organizarão rotinas, operarão negócios e ficarão responsáveis pelos principais serviços públicos. A humanidade passará a viver em um mundo de impressoras de DNA, computadores quânticos, patógenos artificialmente criados, armas autônomas, assistentes robôs e energia abundante. Mas ninguém está preparado.

Em “A próxima onda”, Suleyman e Bhaskar mergulham nas implicações éticas e sociais do uso da IA, abordando temas como privacidade, segurança de dados e viés algorítmico. Na saúde e ciência, ele descreve como a IA revoluciona o diagnóstico médico e a descoberta de medicamentos. O autor também explora o papel da IA no combate às mudanças climáticas e na gestão ambiental, destacando sua utilidade na modelagem climática e conservação de recursos.

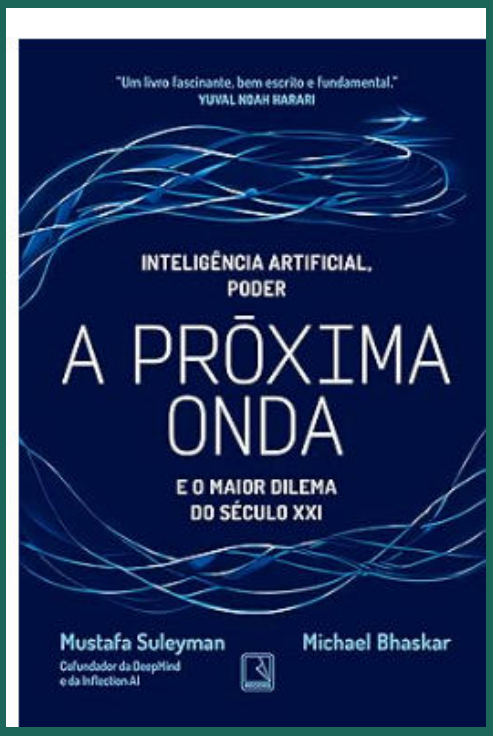



Figura 25 - Capa do livro - A Próxima Onda.

No vídeo mostrado a seguir, "Mustafa Suleyman: *The AI Pioneer Reveals the Future in 'The Coming Wave'* | Intelligence Squared," ele explica como a IA está se tornando uma força motriz, transformando setores e alterando a forma como vivemos e trabalhamos. Ele destaca a importância de compreender e gerenciar essa tecnologia emergente para maximizar seus benefícios e mitigar riscos potenciais.

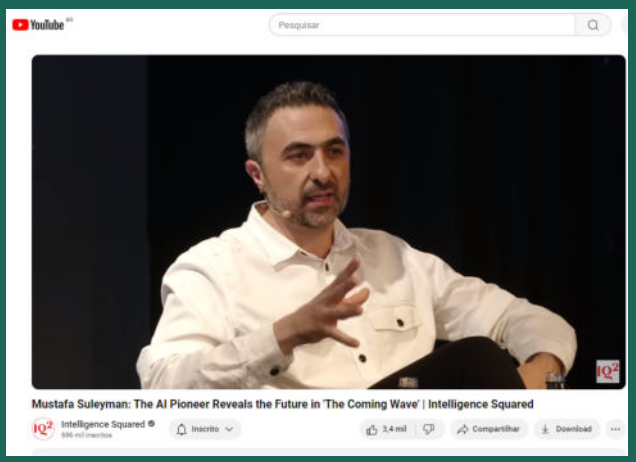



Figura 26- Entrevista de Mustafá Suleyman
https://youtu.be/eJf6QPN9yic

Suleyman explica os avanços recentes e como eles estão sendo aplicados em diferentes áreas, como saúde, transporte e segurança. Uma parte significativa da palestra é dedicada aos desafios éticos e sociais que acompanham a proliferação da IA. Suleyman enfatiza a necessidade de regulamentação e supervisão para garantir que a IA seja desenvolvida e utilizada de forma responsável. Ele aborda questões como viés algorítmico, privacidade de dados e a concentração de poder nas mãos de poucas empresas tecnológicas. Suleyman argumenta que, sem uma abordagem ética e inclusiva, a IA pode exacerbar desigualdades e criar novos problemas sociais. Outro ponto abordado por Suleyman é a transformação do mercado de trabalho. Ele reconhece que a IA substituirá muitos empregos atuais, mas também criará novas oportunidades.

Suleyman enfatiza a importância da requalificação da força de trabalho e da adaptação das políticas educacionais para preparar as pessoas para as novas demandas do mercado. Ele acredita que a educação deve se concentrar em habilidades criativas e de resolução de problemas, que serão cada vez mais valorizadas.

Se você quiser conhecer outros livros sobre a história da IA disponibilizamos um resumo no link: https://aneoescola.wordpress.com/livros/.

Realizada essa introdução vamos apresentar a seguir algumas definições.

A sigla "GPT" significa "*Generative Pretrained Transformer*". Este modelo foi treinado previamente com uma grande quantidade de textos de livros e da internet para lidar com sequências de dados. Isso permite que o programa gere argumentações coerentes sobre muitos tópicos.

Mas, na atualidade, o software tem muitas limitações podendo gerar informações incorretas e também enviesadas. O ChatGPT compreende e produz textos na linguagem natural. Quando você faz uma pergunta, o programa utiliza algoritmos para formular uma resposta adequada. Assim, pode responder sobre diversos assuntos, ajudar a escrever textos, traduzir idiomas e até criar histórias. O algoritmo foi preparado para melhorar continuamente à medida que interage com mais pessoas.

O artigo "*Attention Is All You Need*" é considerado um marco na área de aprendizado de máquina e processamento de linguagem natural, principalmente devido à introdução da arquitetura *Transformer*. Publicado em 2017 por pesquisadores do Google e da Universidade de Toronto, o trabalho revolucionou a forma como modelos de tradução automática e outras tarefas de sequência são abordadas.

# Attention Is All You Need

**Ashish Vaswani*** Google Brain avaswani@google.com

**Noam Shazeer*** Google Brain noam@google.com

**Niki Parmar*** Google Research nikip@google.com

**Jakob Uszkoreit*** Google Research usz@google.com

**Llion Jones*** Google Research llion@google.com

**Aidan N. Gomez*** † University of Toronto aidan@cs.toronto.edu

**Łukasz Kaiser*** Google Brain lukaszkaiser@google.com

**Illia Polosukhin*** ‡ illia.polosukhin@gmail.com

## Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

Figura 27 - Extrato do artigo

“Attention is all you need”

https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Para uma explicação mais acessível: imagine que, ao ler uma frase, você pode focar diretamente nas partes mais importantes para entender o significado, sem precisar ler palavra por palavra na ordem exata.

O Transformer faz algo semelhante, mas em vez de ler uma sequência inteira de dados na ordem, ele usa "atenção" para identificar e se concentrar nas partes mais relevantes da sequência, tudo ao mesmo tempo. Isso permite que o modelo seja muito mais rápido e eficiente, além de obter resultados melhores em tarefas como tradução de idiomas. Essa inovação mudou a maneira como modelos de inteligência artificial processam a linguagem, tornando-os mais eficazes e rápidos. É importante compreender que, embora o ChatGPT pareça inteligente, o programa é essencialmente uma ferramenta probabilística que combina tokens e parâmetros para gerar respostas.

Um token é uma pequena parte de informação que os computadores usam para entender e processar textos. Quando você escreve uma frase ou um parágrafo e um sistema de inteligência artificial vai analisar isso, ele divide o texto em várias partes menores, e essas partes são os tokens.

O algoritmo não pensa ou entende como um ser humano. Entender essa lógica é essencial para gerenciar expectativas demasiadas em relação aos resultados apresentados pelo algoritmo.

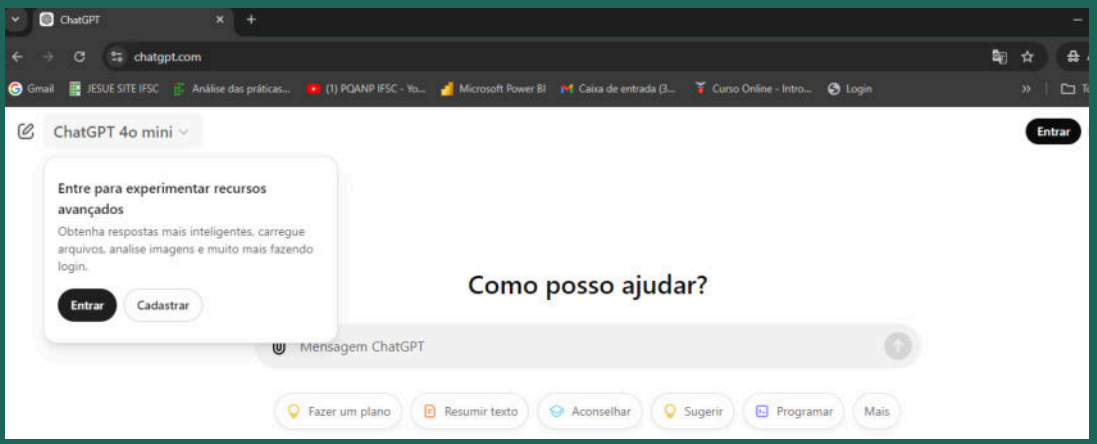



Figura 28- Tela de abertura do ChatGPT

https://chatgpt.com/

Um assunto recorrente na área de Inteligência Artificial são as redes neurais. Uma rede neural artificial, no contexto da inteligência artificial (IA), é uma estrutura computacional inspirada na forma como o cérebro humano processa informações. Ela é composta por unidades interconectadas chamadas neurônios ou nós, organizados em camadas. Essas redes são usadas para modelar e resolver problemas complexos de reconhecimento de padrões, classificação, regressão e outras tarefas que envolvem grandes volumes de dados. A estrutura básica de uma rede neural inclui:

Camada de entrada: Onde os dados iniciais são recebidos. Cada neurônio nesta camada representa uma característica ou variável dos dados de entrada.

Camadas ocultas: São as camadas intermediárias onde ocorre a maior parte do processamento. Neurônios dessas camadas aplicam funções matemáticas sobre os dados, ajustando pesos (valores numéricos) e utilizando uma função de ativação para introduzir não-linearidade, permitindo que a rede capture padrões complexos.

Camada de saída: Onde os resultados são produzidos. Dependendo do problema, pode gerar uma saída que representa uma classificação (por exemplo, identificar uma imagem como "gato" ou "cachorro") ou valores contínuos (como previsões numéricas).

Os neurônios em uma rede neural artificial são interligados por conexões ponderadas.

Durante o treinamento da rede, esses pesos são ajustados usando um processo chamado backpropagation, em conjunto com um algoritmo de otimização como o gradiente descendente. O objetivo é minimizar o erro entre a saída prevista pela rede e o valor real esperado (objetivo do aprendizado supervisionado).Há diferentes tipos de redes neurais, cada uma apropriada para tarefas específicas. Exemplos incluem:

Redes Neurais Artificiais (ANNs): Simples, usadas para problemas gerais de classificação e regressão.

Redes Neurais Convolucionais (CNNs): Usadas principalmente para reconhecimento de imagens e vídeos, são ótimas em captar padrões espaciais.

Redes Neurais Recorrentes (RNNs): Adequadas para processar dados sequenciais, como séries temporais ou processamento de linguagem natural, já que têm a capacidade de "memorizar" informações de entradas anteriores.

Essas redes são extremamente eficazes quando treinadas com grandes quantidades de dados, tornando-se a base para muitos avanços recentes na IA, como reconhecimento de fala, visão computacional, tradução automática e mais. A complexidade e a versatilidade das redes neurais permitiram que a IA alcançasse níveis de desempenho impressionantes em áreas antes dominadas exclusivamente por humanos.

Apesar de algumas definições serem mais complexas para usuários que não são da área da computação, é possível usar a própria IA para aprender e explicar os conceitos de forma didática.

Para isso criamos um aplicativo GPT chamado: Educação com IA.


Figura 29- Aplicativo GPT Educação com IA.

https://chatgpt.com/g/g-tDKxcAItD-educacao-com-ia

Para concluir essa introdução devemos refletir sobre um dos principais riscos da IA na educação, que é a dependência excessiva da tecnologia e das redes sociais.

Quando nos acostumamos a utilizar IA para todas as tarefas, há o perigo de que habilidades críticas, como o pensamento independente, a resolução de problemas e a criatividade, sejam prejudicados. A IA pode fornecer respostas rápidas e precisas, mas não substitui a necessidade de pensarmos criticamente.

Além disso, a implementação de IA na educação pode ampliar as desigualdades existentes se não for realizada de maneira equitativa. Estudantes em regiões menos favorecidas ou em escolas com menos recursos podem não ter acesso às mesmas tecnologias avançadas que seus pares em ambientes mais privilegiados, criando uma divisão digital onde apenas parte dos estudantes se beneficiam das vantagens da IA.

Para contornar alguns destes problemas Salman khan (criador da Khan Academy) realizou uma apresentação explicando como a IA pode ser programada para atuar como professores particulares.


Figura 30- Apresentação de Salman Khan (Khan Academy)

https://youtu.be/hJP5GqnTrNo?si=R-zipVHQflELtHAV

O Khanmigo é uma IA que está ajudando estudantes em diversas disciplinas, desde matemática até programação, sem fornecer respostas diretas, mas guiando o pensamento crítico.

Apesar de entendermos que os professores são insubstituíveis, podemos imaginar que em uma ou mais décadas a IA poderá ocupar um papel mais relevante nas salas de aula.

Embora a IA Generativa tenha um grande potencial para contribuir para a elevação do aprendizado, é essencial abordar os riscos associados. Governos, instituições e desenvolvedores devem trabalhar juntos para criar diretrizes e políticas que garantam uma implementação ética. Assim poderemos maximizar os benefícios da IA, ao mesmo tempo em que mitigamos seus riscos.

Um outro risco significativo da IA é o viés e a discriminação. Apesar dos esforços das grandes empresas em reduzir esse problema, a IA é tão imparcial quanto os dados com os quais o algoritmo foi treinado.

Realizamos um experimento para avaliar esse problema. Vejam essa imagem a seguir:

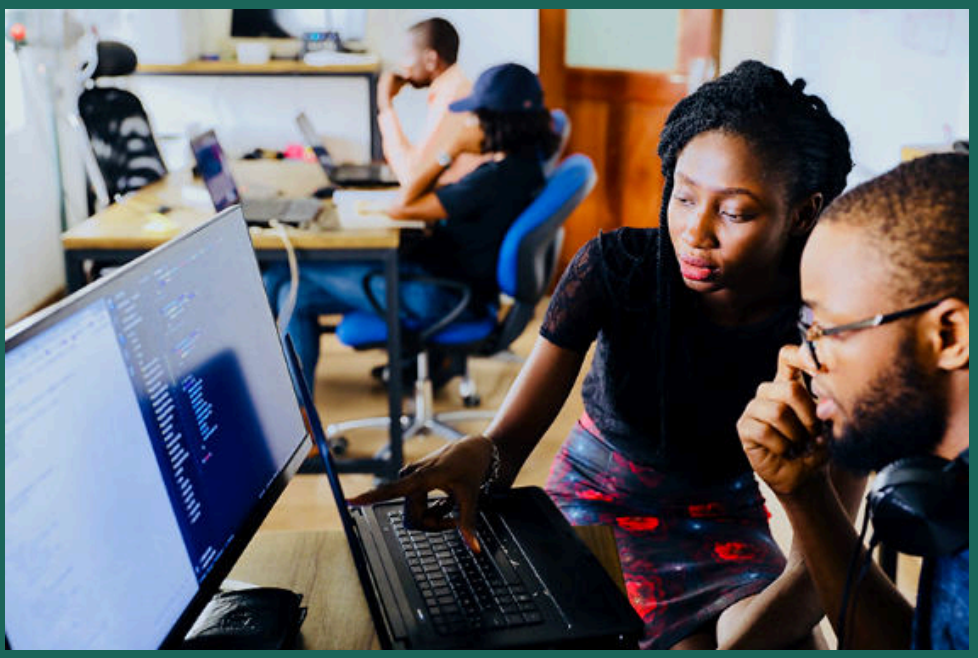


Figura 31- Imagem real de dois jovens

https://portalmulheramazonica.com.br/site/noticia/carrefour-tera-de-pagar-mais-de-800-bolsas-para-estudantes-negros/

Acessamos o aplicativo de prompt GPT https://chatgpt.com/g/g-UpAtTm6O8-rei-do-prompt e solicitamos que o GPT transformasse a foto em um texto descritivo. “Crie uma descrição muito detalhada da imagem”. Como resultado obtivemos:

“A imagem mostra um ambiente de trabalho colaborativo, com foco em dois indivíduos que parecem estar em um processo de orientação ou discussão técnica. À esquerda, uma mulher está sentada em frente a um laptop aberto, apontando diretamente para a tela com a mão direita, enquanto explica algo ao homem à direita, que está observando atentamente e segurando o queixo com a mão, em sinal de concentração. Ambos estão sentados lado a lado em cadeiras de escritório. O laptop exibe uma interface de desenvolvimento de software, com linhas de código visíveis, sugerindo que a atividade principal gira em torno de programação ou revisão de código. O homem usa um headset, o que pode indicar que ele está envolvido em um ambiente de trabalho técnico que exige comunicação frequente, talvez com outros membros da equipe ou clientes. A mulher está vestida com uma camisa preta e uma saia com

detalhes vermelhos, enquanto o homem veste uma camiseta preta. Ao fundo, desfocados, vemos outros dois indivíduos também envolvidos em tarefas semelhantes. Eles estão sentados em cadeiras de escritório e trabalham em seus respectivos computadores, criando a impressão de que todos estão inseridos em um ambiente de escritório compartilhado, como um espaço de coworking ou o setor de tecnologia de uma empresa. O ambiente parece ser bem iluminado, com uma atmosfera de produtividade. A interação entre a mulher e o homem na frente parece central para a imagem, com o gesto de apontar para a tela sugerindo que ela está fornecendo orientação ou instrução, enquanto ele absorve as informações. Isso indica uma dinâmica de aprendizado ou colaboração, possivelmente relacionada ao desenvolvimento de software ou à execução de um projeto técnico.”

Na sequência, abrimos uma nova aba e solicitamos ao ChatGPT que criasse uma imagem a partir da descrição. O resultado é mostrado na sequência. As imagens trazem estereótipos de dois jovens brancos em um espaço acadêmico.

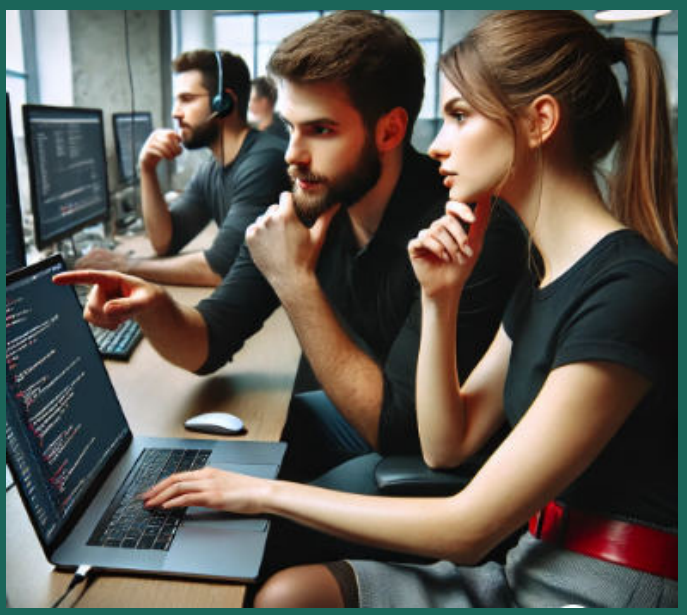


Figura 32- Imagem construída pelo ChatGPT

Após repetir esse procedimento algumas vezes foi possível construir prompts mais precisos.

Figura 33 - Imagem reconstruída pelo ChatGPT

Na imagem acima corrigimos parte do problema dos padrões do algoritmo solicitando que a professora seja mais velha e que o jovem seja negro. Mas isso vamos explorar no Minicurso II.

# AUTOAVALIAÇÃO

**1- Qual foi o principal objetivo da criação do teste proposto por Alan Turing ?**

A) Avaliar a capacidade das máquinas de realizar cálculos matemáticos complexos.

B) Determinar se uma máquina pode imitar o intelecto humano ao ponto de ser indistinguível de um humano.

C) Desenvolver programas de computador para jogar xadrez.

D) Identificar a melhor forma de armazenar grandes quantidades de dados.

E) Criar sistemas de segurança para proteger dados sensíveis.

**2- Qual é uma das principais diferenças entre aprendizado de máquina (machine learning) e deep learning mencionadas no curso?**

A) O aprendizado de máquina não utiliza algoritmos.

B) O deep learning utiliza redes neurais profundas, enquanto o aprendizado de máquina não.

C) O aprendizado de máquina é uma técnica mais recente que o deep learning.

D) O deep learning é aplicado exclusivamente em sistemas de navegação.

E) O aprendizado de máquina não tem aplicações práticas.

**3- Quais são alguns dos impactos sociais da automação e da IA no mercado de trabalho, conforme discutido no curso?**

A) Aumento de salários em todas as profissões.

B) Criação de empregos exclusivamente na área de TI.

C) Potencial perda de empregos devido à automação e a necessidade de adaptação educacional.

D) Exclusão das tecnologias de IA em setores financeiros.

E) Redução das horas de trabalho para todos os funcionários.

**4- No contexto do curso, como a IA pode melhorar a experiência de aprendizado em ambientes educacionais?**

A) Removendo a necessidade de professores humanos.

B) Personalizando o ensino de acordo com as necessidades individuais dos alunos.

C) Substituindo todos os métodos tradicionais de ensino.

D) Automatizando todas as atividades extracurriculares dos alunos.

E) Reduzindo a interação entre alunos e professores.

Confira suas respostas:

| | | | |
|---|---|---|---|
| 1B | 2B | 3C | 4B |

**RESUMO:**

Este capítulo introduziu a história e as principais definições da Inteligência Artificial (IA), explorando sua evolução desde os primeiros conceitos até as tecnologias avançadas que temos hoje. O objetivo é fornecer uma compreensão sólida das origens da IA, seus marcos históricos, e as definições que norteiam este campo de estudo.

A jornada da inteligência artificial começou nos anos 1950, quando o matemático Alan Turing questionou se as máquinas poderiam pensar. Esse questionamento levou à criação do teste de Turing, um critério para determinar se uma máquina pode imitar o intelecto humano: “se um humano interagir com um computador e não conseguir distinguir se está interagindo com um humano ou uma máquina, podemos considerar essa máquina inteligente”.

John McCarthy, Marvin Minsky e outros pesquisadores, em uma conferência em 1956, lançaram as bases para o desenvolvimento da inteligência artificial. A definição inicial de IA era a de um campo de estudo dedicado ao desenvolvimento de máquinas que pudessem realizar tarefas que normalmente requerem inteligência humana, como aprender, raciocinar e resolver problemas.

A história da IA é marcada por períodos de grande entusiasmo e subsequente desilusão, conhecidos como "AI Winters". O

primeiro desses períodos ocorreu nos anos 1970, devido às limitações da tecnologia da época e ao otimismo exagerado que levou a promessas não cumpridas. Um segundo AI Winter ocorreu nos anos 1980 após o declínio do interesse e financiamento em tecnologias de IA.

Nos anos 1990 e início dos anos 2000, o interesse pela IA renasceu, impulsionado pelos avanços em algoritmos de aprendizado de máquina e o aumento da capacidade computacional. Esse período também viu o surgimento do deep learning, que revolucionou campos como o processamento de linguagem natural e a visão computacional, graças a redes neurais profundas.

Atualmente, a IA permeia muitos aspectos do dia a dia. Assistentes virtuais como Siri e Alexa, recomendações personalizadas em serviços de streaming como Netflix e Spotify, e sistemas de navegação em tempo real utilizam IA para melhorar a experiência do usuário. Além do uso cotidiano, a IA tem aplicações transformadoras em várias indústrias. Na saúde, algoritmos de IA são usados para diagnóstico de imagens médicas com precisão comparável ou superior à dos humanos. No setor financeiro, a IA ajuda na detecção de fraudes e na personalização de serviços financeiros. Na educação, sistemas adaptativos de aprendizagem utilizam IA

para personalizar o ensino de acordo com as necessidades individuais dos alunos.

A representação da Inteligência Artificial (IA) como uma invenção apocalíptica é comum na cultura popular, especialmente no cinema. Alguns filmes mostram a IA como uma força perigosa, onde robôs se tornam incontroláveis e colocam a humanidade em risco.

Por outro lado, há filmes que retratam a IA como benéfica, ajudando a resolver problemas complexos e melhorando a vida humana. Filmes como "2001, uma Odisséia no Espaço", "Blade Runner", "Eu, Robô", "Ex Machina" e "Her" exploram diferentes aspectos e possibilidades da IA.

As aplicações de IA também levantam questões importantes sobre impacto social, incluindo preocupações com privacidade, segurança e a potencial perda de empregos devido à automação. É crucial que os educadores estejam equipados para orientar seus alunos através destas questões complexas e em constante evolução.

Pesquisadores como Nick Bostrom, em "Superinteligência", e Kai-Fu Lee, em "Inteligência Artificial", discutem as implicações éticas e existenciais dessa transição, destacando os riscos de uma IA descontrolada e as estratégias necessárias para garantir seu uso seguro.

# AVALIAÇÃO DO MINICURSO I

1- Utilize o GPT: ARTICLE ANALYST e faça uma análise crítica do artigo: https://nickbostrom.com/ethics/ai.pdf

LINK DO APLICATIVO:

https://chatgpt.com/g/g-JwYhftH3M-article-analyst

2- Utilize o GPT: RESUME AI e faça um resumo da entrevista de Ronaldo Lemos: https://youtu.be/UYgyvMBZYSw

LINK DO APLICATIVO:

https://chatgpt.com/g/g-g6EQeCbix-resume-ai

3- Utilize o GPT: RESUME AI e faça um resumo do vídeo:

https://youtu.be/95kwgiOO9mA

Envie essa tarefa em PDF para o email: iacursobasicointeligenciaartif@gmail.com