

A close-up photograph of a person's hand holding a glowing incandescent lightbulb. The bulb is illuminated from within, casting a warm glow. The hand is positioned palm-up, with the lightbulb resting on the palm. In the background, there are some dark, out-of-focus objects, possibly books or papers, suggesting a workspace or study area.

A evolução da Inteligência Artificial: abrindo a caixa preta.

Fabrício Barth

Fabrício Barth

Formado em Ciência da Computação e com Doutorado em Engenharia Elétrica pela Universidade de São Paulo. Desenvolve projetos relacionados com Aprendizagem de Máquina e Inteligência Artificial em diversas áreas (i.e., financeira, Internet, segurança pública, educação e recursos humanos) desde 2003. Atualmente ocupa a posição de professor no Insper.

<https://www.linkedin.com/in/fbarth/>
fabriciojb@insper.edu.br

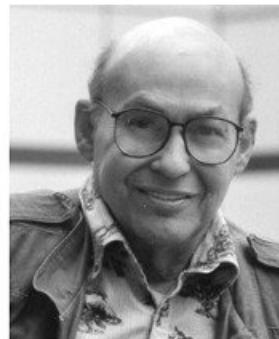


- Criação da área de Inteligência Artificial (IA).
- Solucionadores gerais de problemas.
- A relação da inteligência humana com a inteligência da máquina.

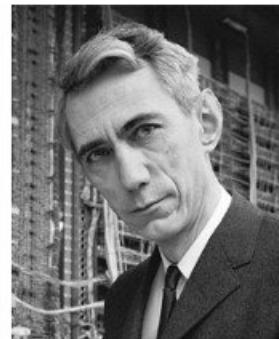
1956 Dartmouth Conference: The Founding Fathers of AI



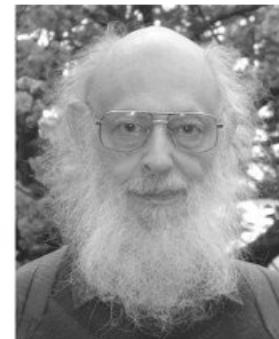
John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



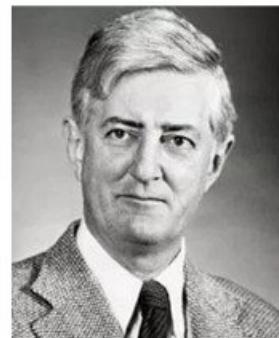
Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

Muitos projetos relacionados com IA foram desenvolvidos de 1950 até 2015. Mas a grande maioria no campo acadêmico.

No entanto, criações relacionadas com automação ou computação sempre impactaram as nossas vidas.

Imagen de 1988

Page 10A The Daily Item — Sumter, S.C. Saturday, April 5, 1986

AP photo

Elementary school teachers picket against use of calculators in grade school
The teachers feel if students use calculators too early, they won't learn math concepts

Math teachers protest against calculator use

By JILL LAWRENCE

"My older kids don't pay any strate," he said. "Teachers are attention to an answer being absurd."

DeepBlue vs Kasparov (Xadrez)

1997: primeira vez que uma máquina ganha uma partida de Xadrez contra um campeão mundial.

1950: Claude Shannon. Primeiro a propor o desenvolvimento de um agente que joga Xadrez usando Min-Max.

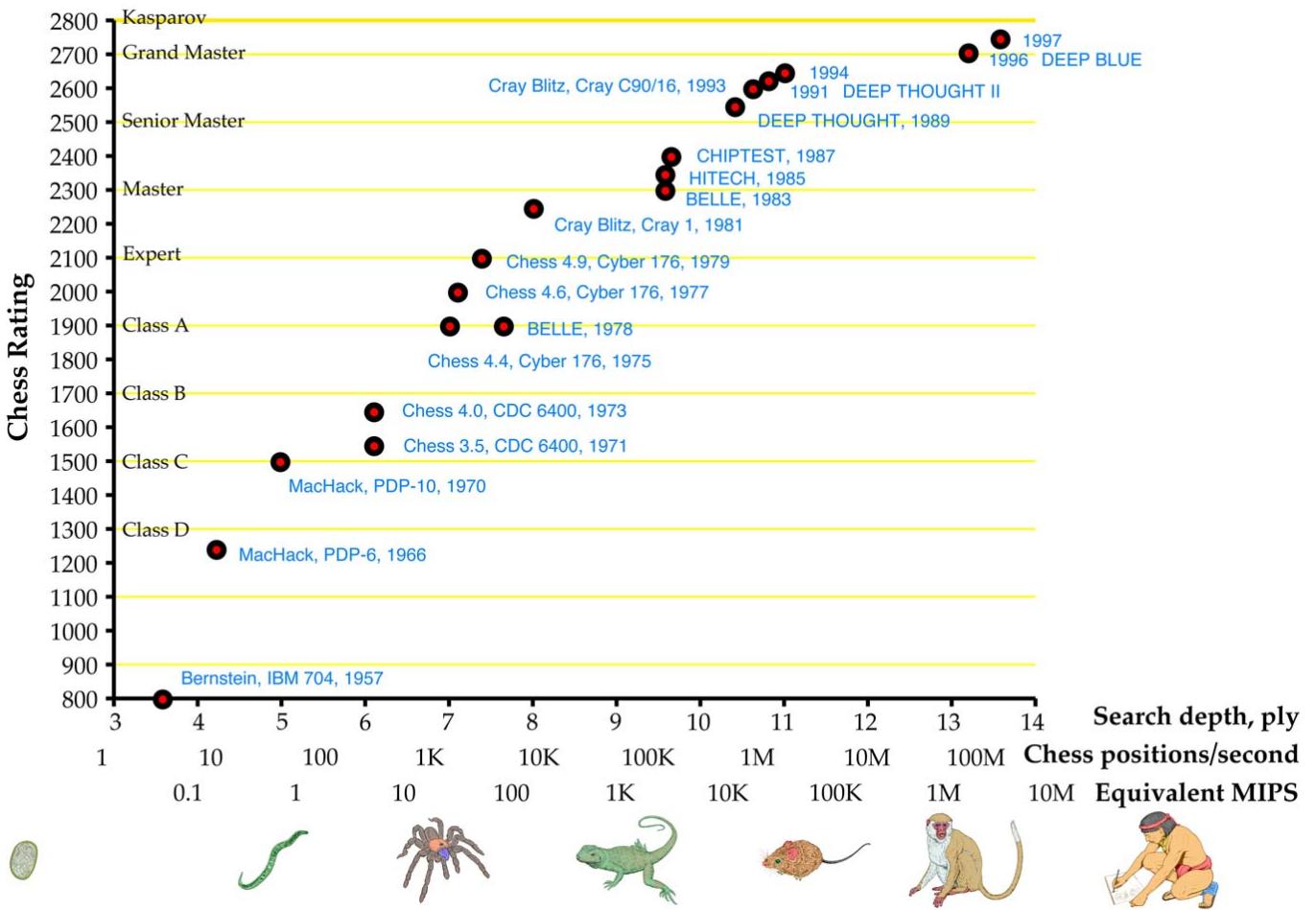


**KASPAROV
AND
DEEP BLUE**

O que mudou de Shannon (1950) até o Deep Blue (1997)?

- Poder computacional!
- Basicamente, todas as soluções propostas ao longo do tempo implementam a mesma ideia.

Chess Machine Performance versus Processing Power

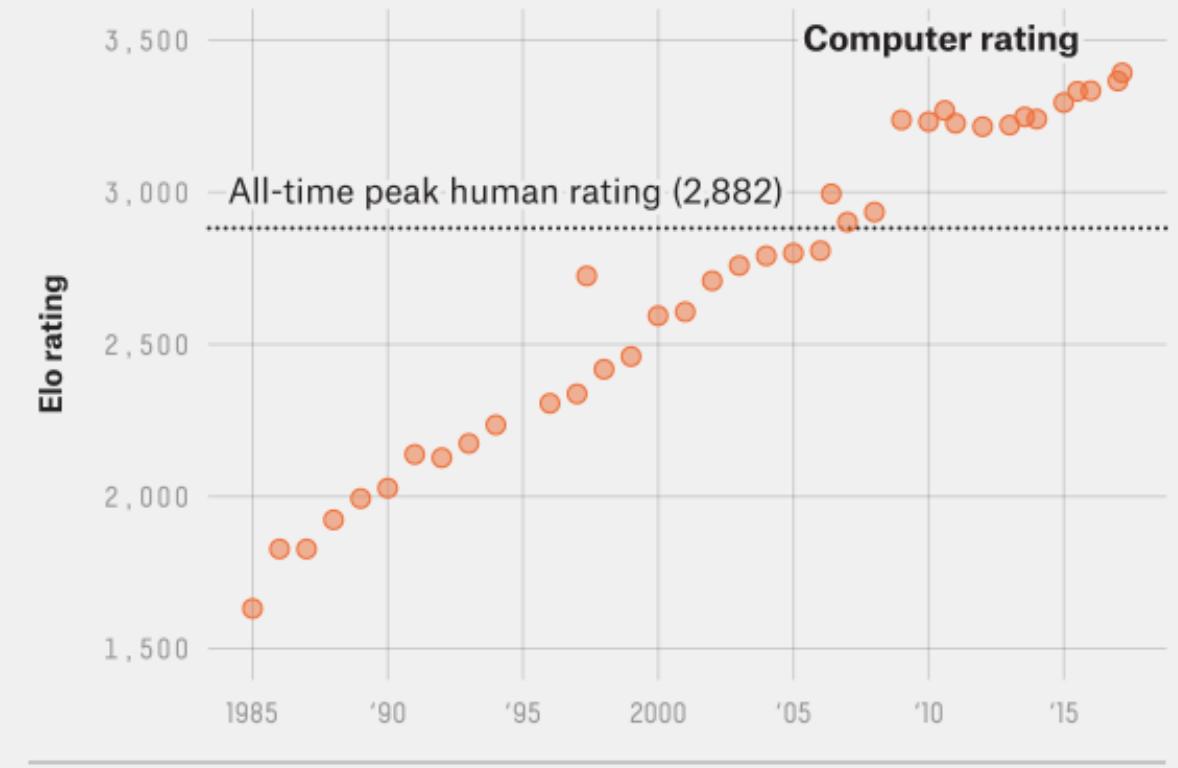


Agora não tem mais graça jogar Xadrez (!?)

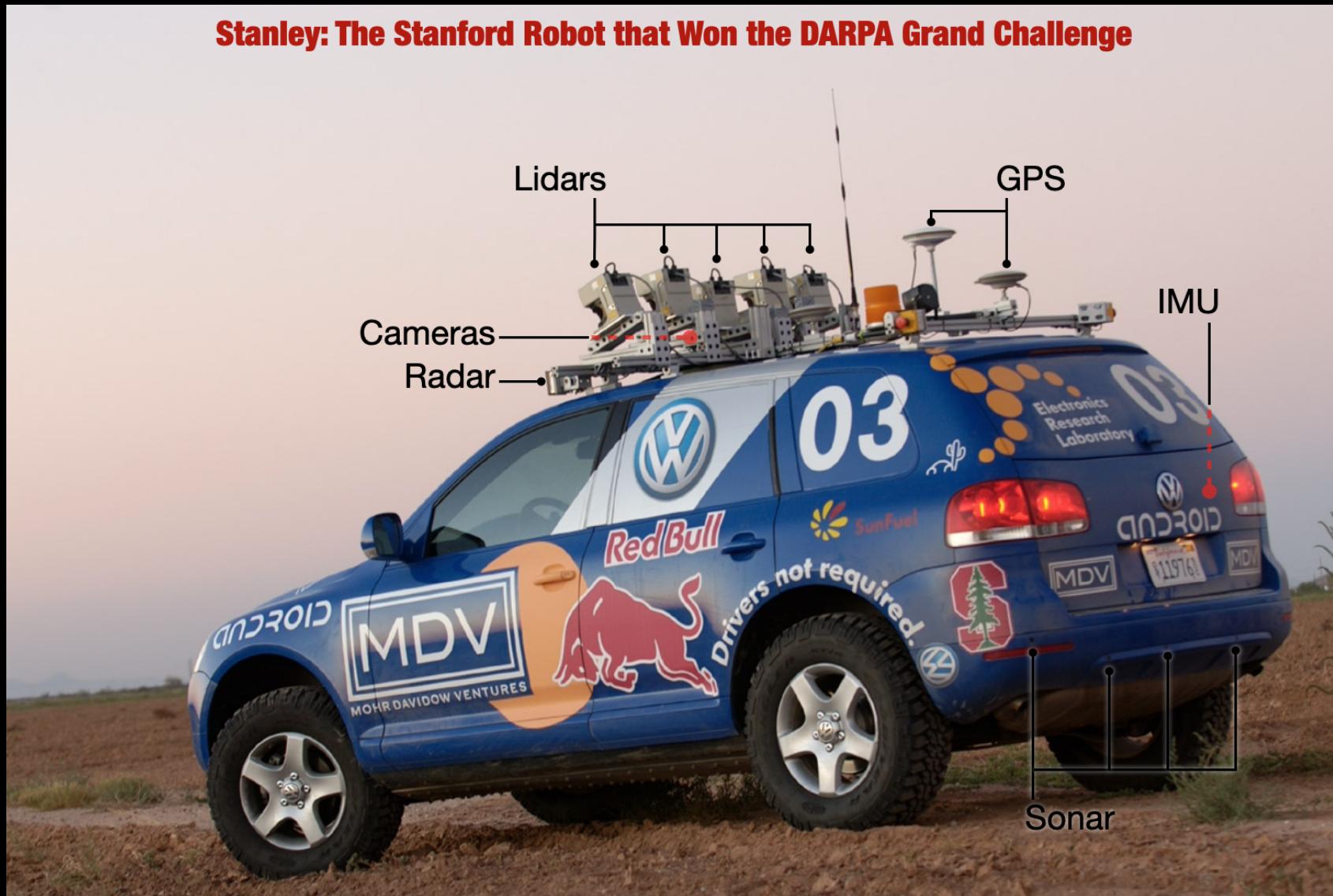
Xadrez nos dias atuais

- Xadrez não deixou de existir.
Inclusive, o número de jogadores aumentou.
- O computador se tornou a **principal** ferramenta de treinamento para os jogadores.
- <https://stockfishchess.org/>

Chess computers have been better than humans for years
Elo ratings since 1984



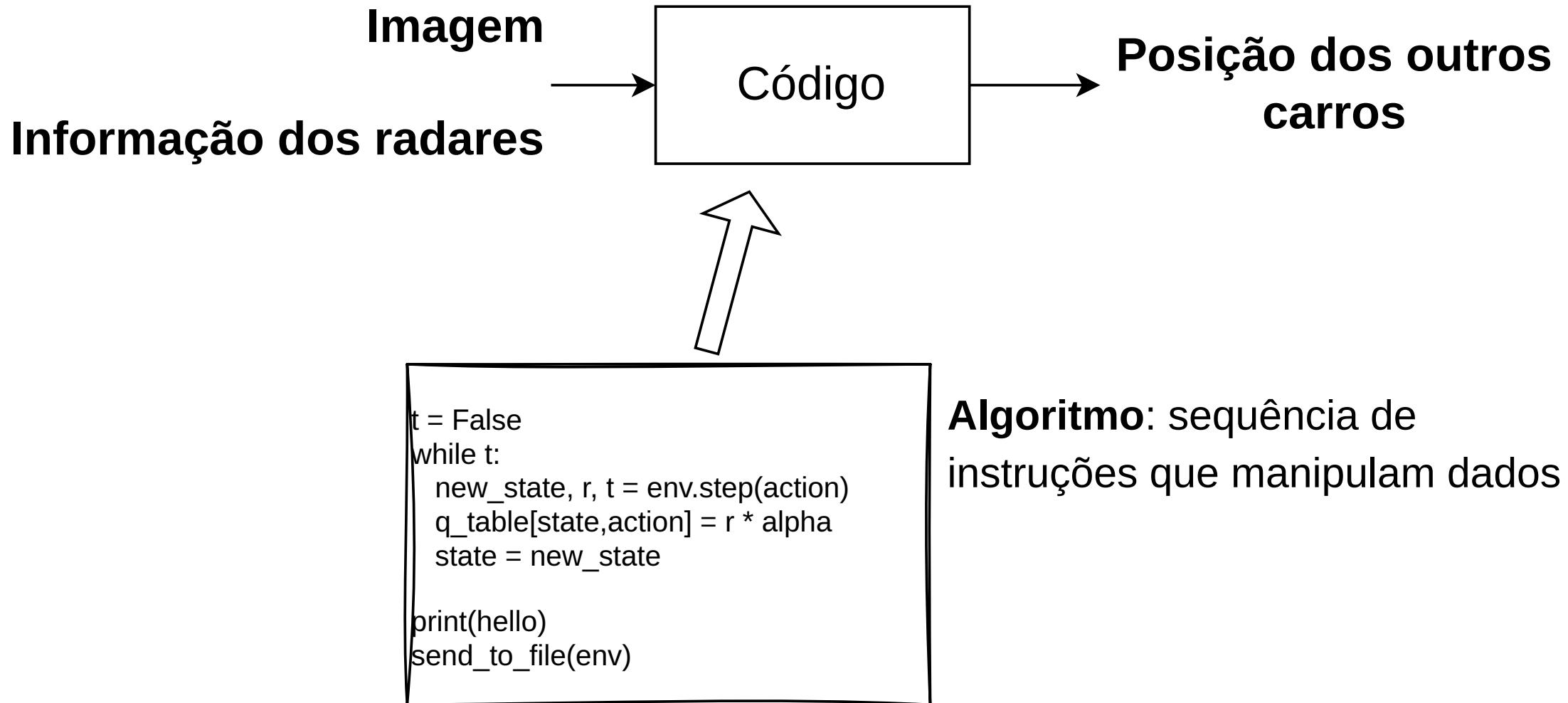
Veículos Autônomos: DARPA Challenge (2005)



Identificação de outros carros na rua



Identificação de outros carros na rua



Aprendizagem de Máquina Supervisionada

Exemplos

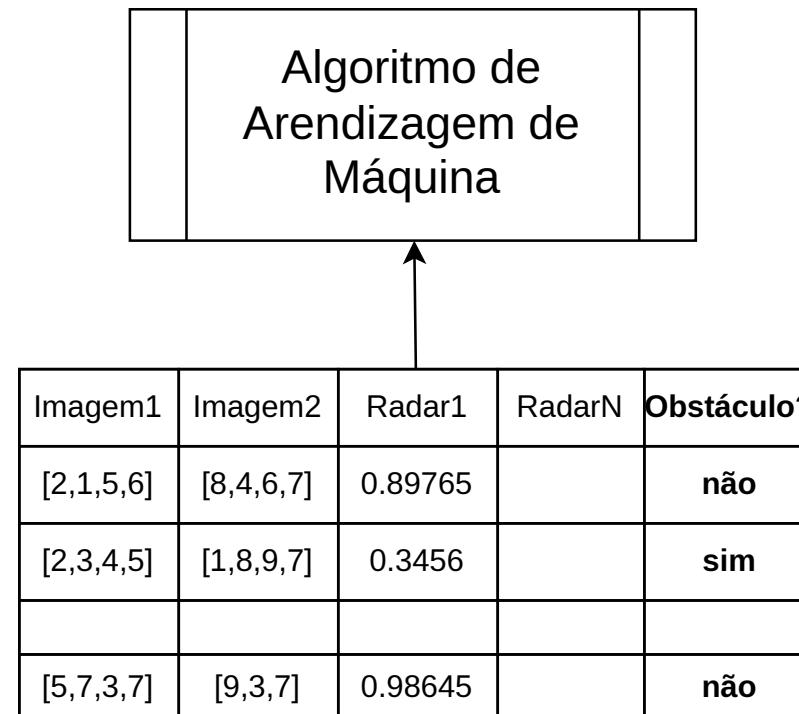
Atributos

Atributo especial

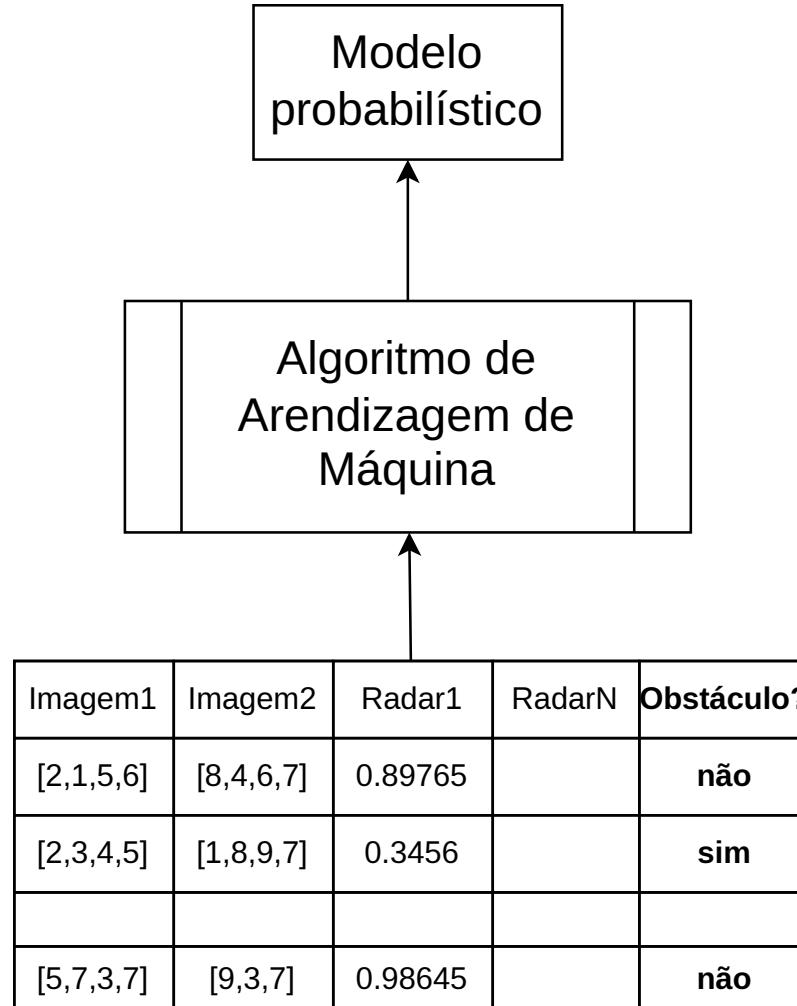
The diagram illustrates a supervised learning dataset. On the left, a brace groups five rows of data under the label 'Exemplos'. Above the first four rows, another brace groups them under the label 'Atributos'. An arrow points from the label 'Atributo especial' to the column header 'Obstáculo?'.

Imagen1	Imagen2	Radar1	RadarN	Obstáculo?
[2,1,5,6]	[8,4,6,7]	0.89765		não
[2,3,4,5]	[1,8,9,7]	0.3456		sim
[5,7,3,7]	[9,3,7]	0.98645		não

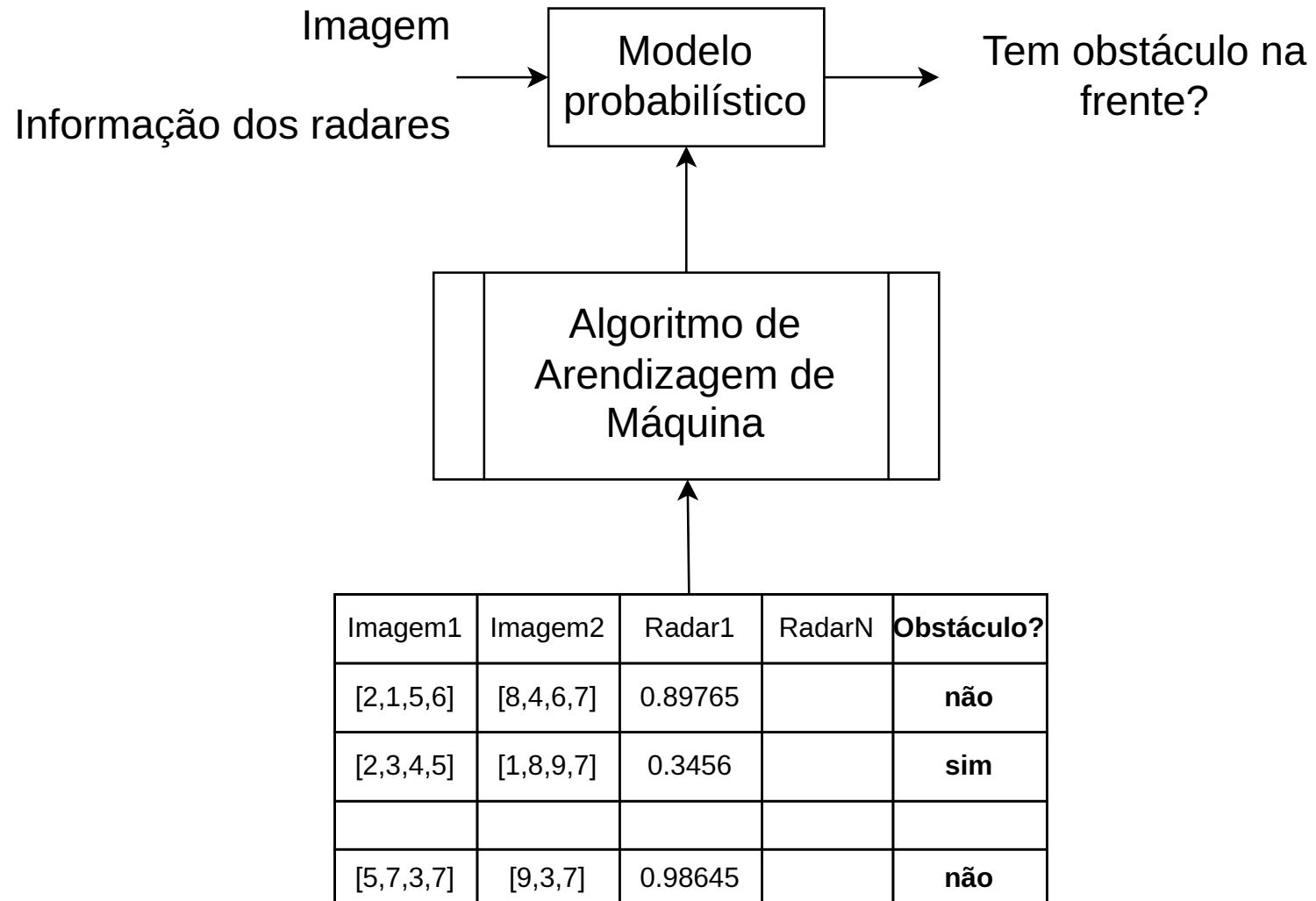
Aprendizagem de Máquina Supervisionada



Aprendizagem de Máquina Supervisionada

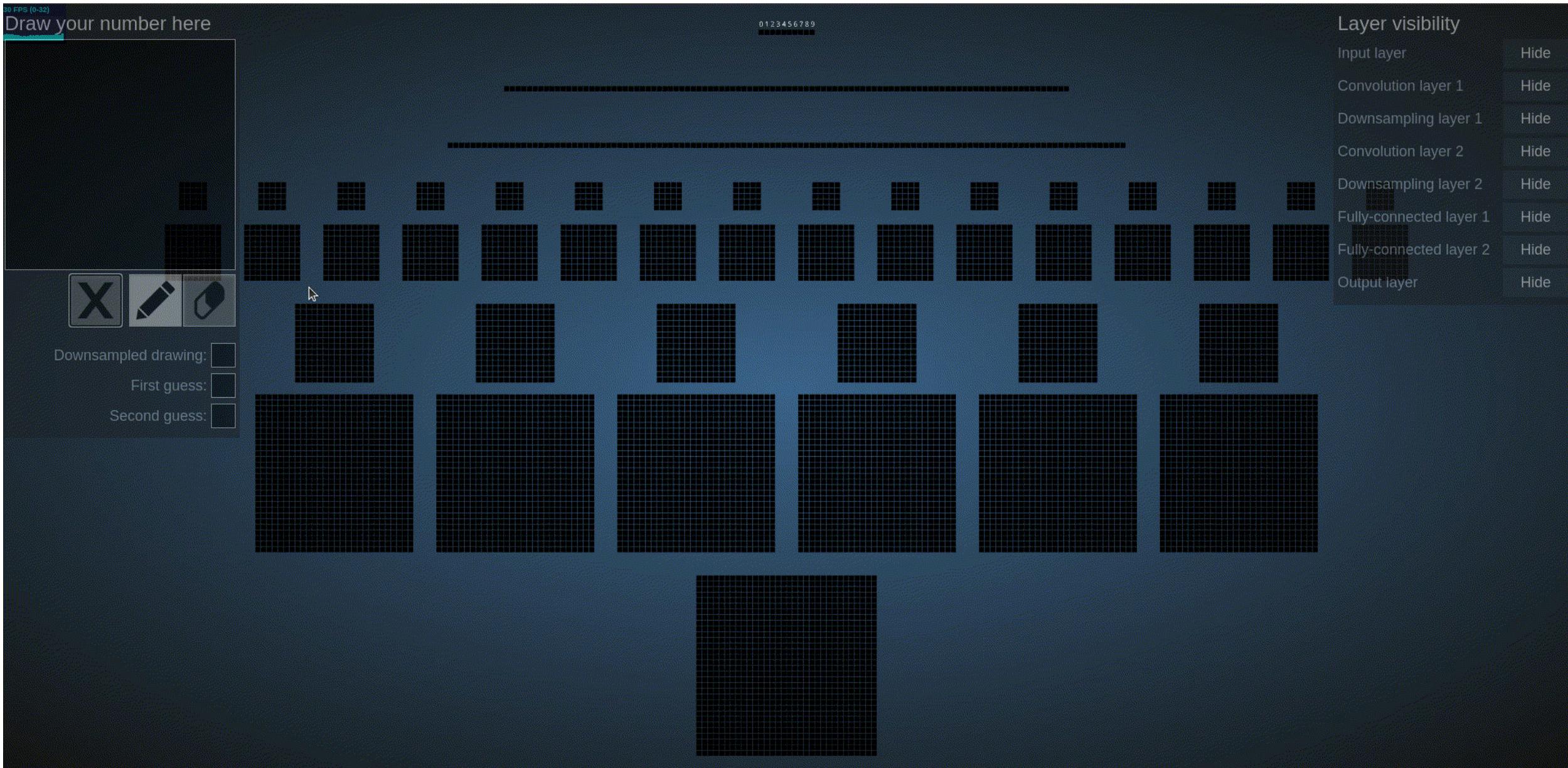


Aprendizagem de Máquina Supervisionada



Aprendizagem de Máquina Supervisionada

- Filtro de spam: classificação de e-mails em spam ou não spam.
 - vários emails rotulados como spam ou não.
 - algoritmo aprende a classificar novos emails.
- Que propaganda ofertar para os meus usuários?
 - histórico de propagandas e cliques dos usuários.
- Controle de qualidade: classificação de frutas em boas ou ruins.
 - histórico de frutas classificadas por humanos.
- Reconhecimento de dígitos manuscritos: classificação de imagens de dígitos manuscritos em 0, 1, 2, ..., 9.



Impactos (2010 → dias atuais)

- Exemplos de aplicações *business-to-consumer (B2C)*:
 - **Assistentes virtuais**: Siri, Google Assistant, e Alexa
 - **Assistentes virtuais para atendimento**: chatbots
 - **Recomendação de conteúdo**: Netflix, Spotify, Youtube
 - **Filtros de spam**: Gmail, Outlook.
 - Waze, Google Maps.
- Exemplos de aplicações *business-to-business (B2B)*:
 - detecção de fraude,
 - análise de crédito,
 - controle de qualidade (classificação de frutas, peças, etc.),
 - definição de rotas (Uber, iFood),
 - **interminável lista de aplicações...**

O que estes exemplos têm em comum?

- São softwares que conseguem executar um **objetivo complexo, muito específico** e de forma **autônoma**.
- No entanto:
 - Os objetivos destes softwares são definidos por humanos.
 - Os dados analisados por estas implementações foram definidos por humanos.
 - Estas implementações **não** tem capacidade de aprender novas tarefas sem a intervenção humana.
- Mas, este tipo de tecnologia **permite automatizar processos que outras tecnologias não permitem**.

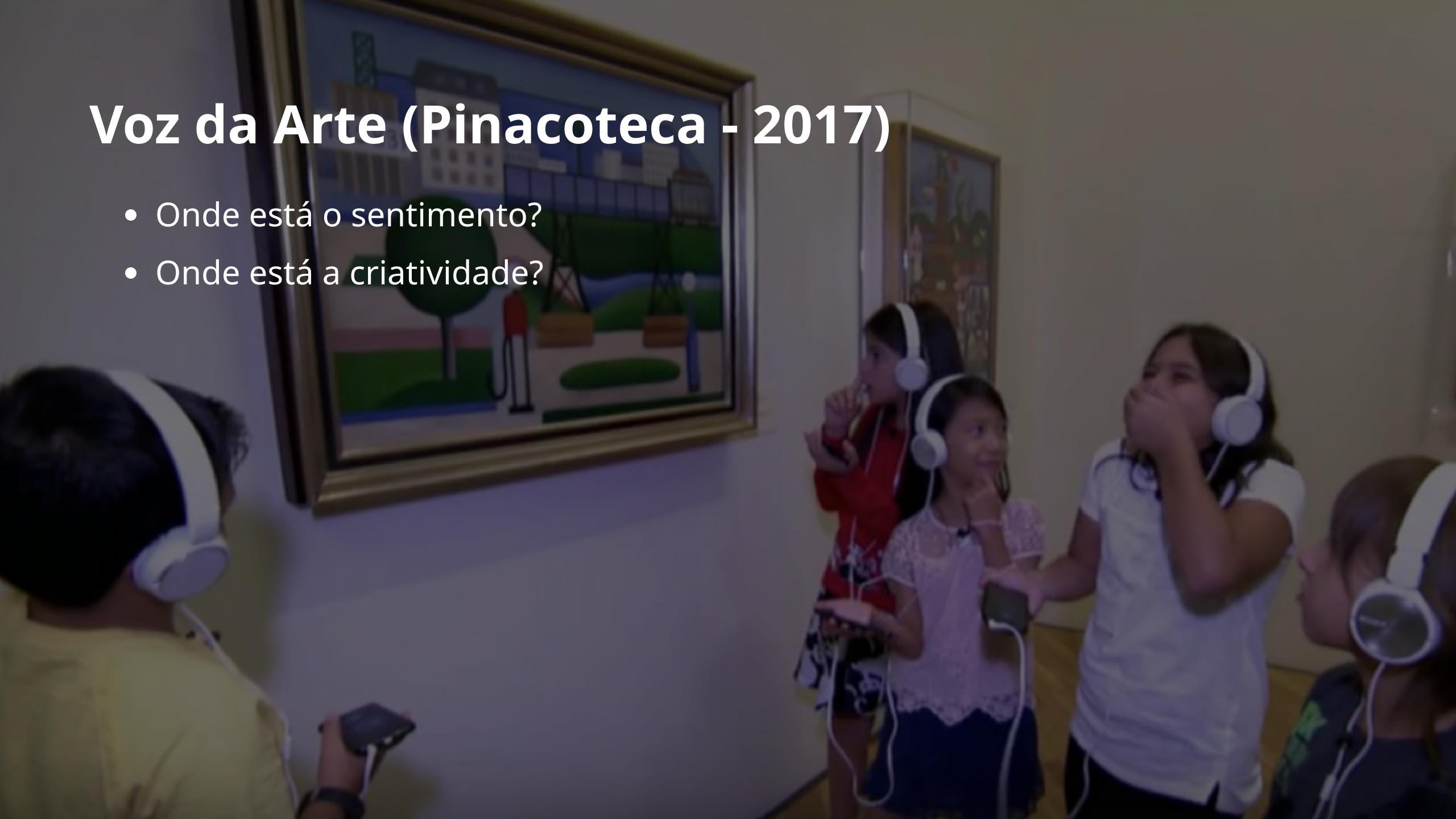
Voz da Arte (Pinacoteca - 2017)



A Voz da Arte - IBM Watson: <https://www.youtube.com/watch?v=1rOAgvCnZpw>

Voz da Arte (Pinacoteca - 2017)

- Onde está o sentimento?
- Onde está a criatividade?



E o ChatGPT?

ChatGPT (2022)

- Trata-se de um modelo probabilístico que estima a probabilidade de uma palavra dado um contexto.
- Exemplo de *Large Language Model* (LLM). Estima-se que este modelo foi treinado usando mais de 45TB de texto ([Common Crawl dataset](#)).
- Imita o comportamento de um humano em uma conversa.

<https://chat.openai.com/>



How can I help you today?

Help me pick
an outfit that will look good on camera

Give me ideas
for what to do with my kids' art

Come up with concepts
for a retro-style arcade game

Brainstorm names
for my fantasy football team with a frog theme

Message ChatGPT...



Podemos confiar no ChatGPT? Não!

- A OpenAI não divulga dados sobre o conjunto de treinamento, nem sobre o processo de treinamento do modelo.
- O ChatGPT tem alucinações e faz afirmações falsas.
- Pode fornecer referências falsas.
- mas, mesmo assim pode ser útil:
 - sumarização de texto;
 - tradução automática;
 - geração de código ([Copilot](#));
 - classificação de texto (análise de sentimento)

Mas consegue passar no exame da OAB!

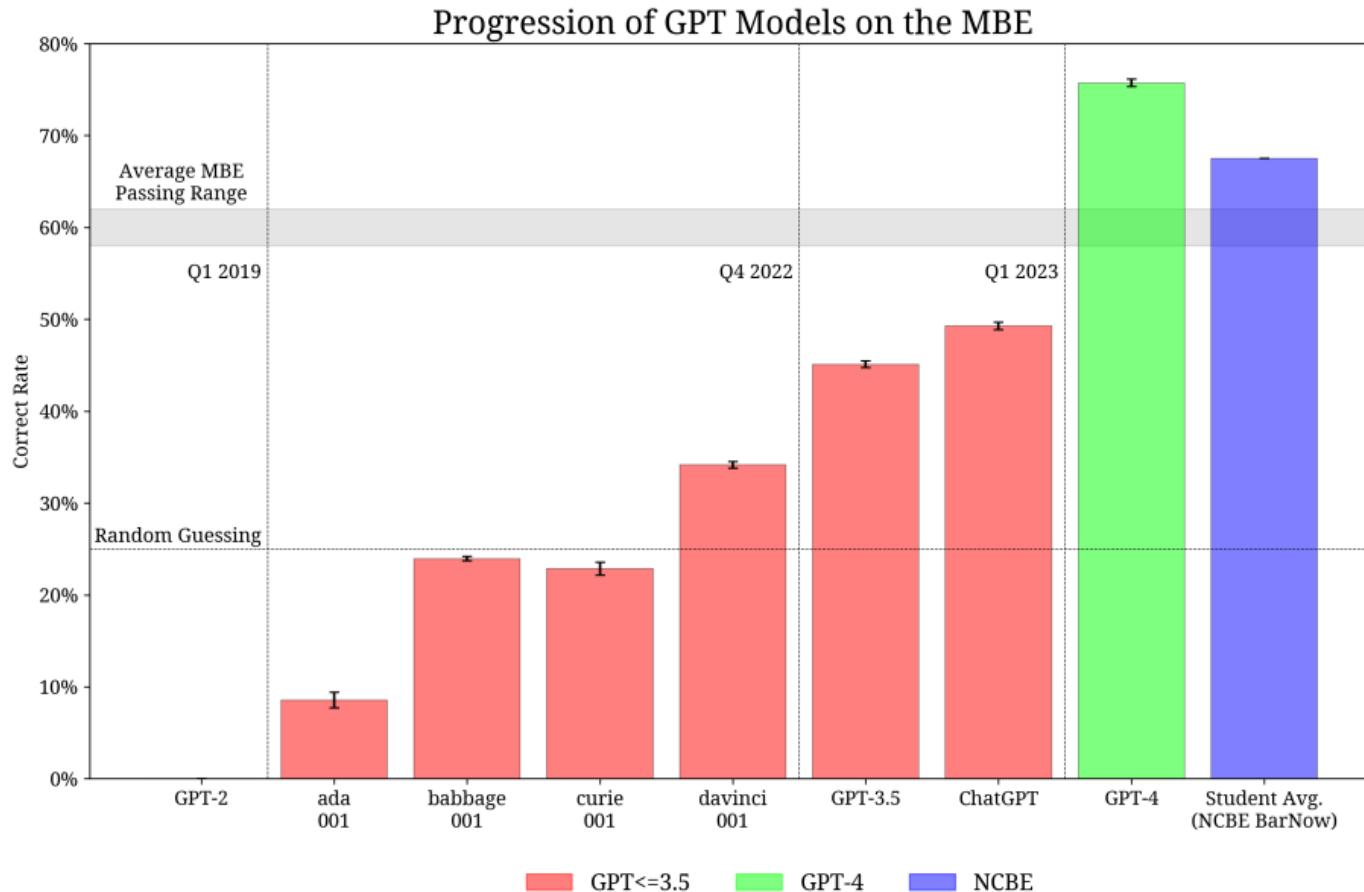


Figure 1. Progression of Recent GPT Models on the Multistate Bar Exam (MBE)

LLMs: próximos passos

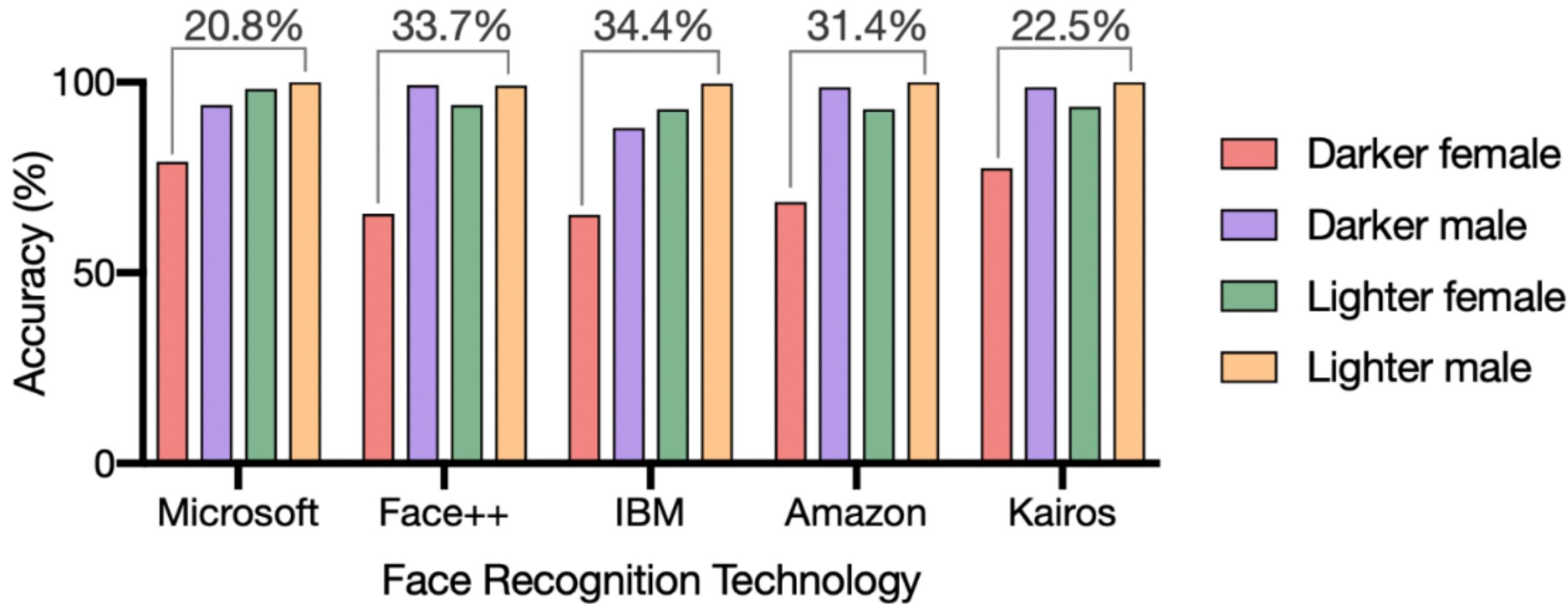
- Tornar os LLMs com comportamento mais factual, coerente e confiável.
- Reduzir a necessidade de dados e poder computacional para o treinamento e inferência. No futuro todos teremos o nosso próprio LLM (Stockfish).
- LLMs especializados para diferentes domínios e usos.
- LLMs multi-modal.

Impactos sociais

Dados, Poder Computacional, Algoritmos e Pessoas permitem o desenvolvimento de novas aplicações de IA em novos domínios.

- **Problema de Privacidade** devido a alta procura por dados gerados pelas pessoas.
- **Problema de Viés** devido a falta de diversidade de dados e pessoas envolvidas no desenvolvimento de soluções de IA.

Accuracy of Face Recognition Technologies



Impactos sociais

Dados, Poder Computacional, Algoritmos e Pessoas permitem o desenvolvimento de novas aplicações de IA em novos domínios.

- **Problema de Privacidade** devido a alta procura por dados gerados pelas pessoas.
- **Problema de Viés** devido a falta de diversidade de dados e pessoas envolvidas no desenvolvimento de soluções de IA.
- **Falta de transparência e interpretabilidade das decisões** tomadas por sistemas de IA.

Considerações finais

- IA é uma área de pesquisa que existe há mais de 70 anos e está em constante evolução **incremental**.
- A maioria das aplicações que hoje estão disponíveis no mercado são baseadas em **aprendizagem de máquina supervisionada**.
- *Nenhum [designer] vai perder o emprego para a IA. Mas, [designers] que não usam IA vão perder o emprego para [designers] que usam IA.*

Referências adicionais

- Game over: Kasparov and The Machine, 2003
- Hans Moravec. When will computer hardware match the human brain?, 1997.
- Chess's New Best Player Is A Fearless, Swashbuckling Algorithm, 2018.
- Erik Brynjolfsson, Tom Mitchell. What can machine learning do? Workforce implications. Science, 2017.
- Erik Brynjolfsson, Tom Mitchell. Track how technology is transforming work. Nature, 2017.