

Introdução à Machine Learning

Visão geral e conceitos introdutórios

Agenda

01

Apresentação
do curso

02

O que é IA,
ML e DL?

03

A importância
dos dados

04

Dilemas e
problemáticas

05

Aprendizado e
tarefas em ML

06

Exercícios de
revisão

01

Apresentação do curso



Um pouco sobre a Iris

Existe **desde 2019**.

Busca promover a comunidade de **IA, Machine Learning e Data Science** na Unicamp.

Os integrantes são da **Computação**, majoritariamente, e de **outros cursos**.

Equipe do curso: por volta de **7 pessoas**.

Temos conteúdo no **nossa canal do YouTube!**



Como será o curso?

Formas de participar: presencial (IC300 / CC00) e online (Google Meet).

Horários: 14h00 até às 17h00 (3 horas de duração).

Canal de comunicação: grupo de WhatsApp.

Aulas gravadas serão disponibilizadas no YouTube da Iris.

As aulas terão partes **expositivas e práticas**.

Pré-requisitos:

- Obrigatórios:
 - Python, NumPy, Pandas e Matplotlib;
 - Álgebra linear básica (operações com matrizes).
- Opcionais:
 - Seaborn;
 - Cálculo (noções de derivada, derivada parcial e vetor gradiente) e outros conceitos de álgebra linear.

Cronograma



Aula 1 (18/05):

Introdução à Machine Learning



Aula 2 (01/06):

Aprendizado superv. e Reg. Linear



Aula 3 (22/06):

Otimização da função de custo



Aula 4 (29/06):

Avaliação de modelos e Reg. Logística



Aula 5 (13/07):

Árvores de decisão e *random forest*



Aula 6 (27/07):

Aprendizado não-superv. e clusterização



Aula 7 (10/08):

Redução de dimensionalidade



Aula 8 (24/08):

Introdução à Redes Neurais

As datas poderão (e provavelmente vão) sofrer alterações.

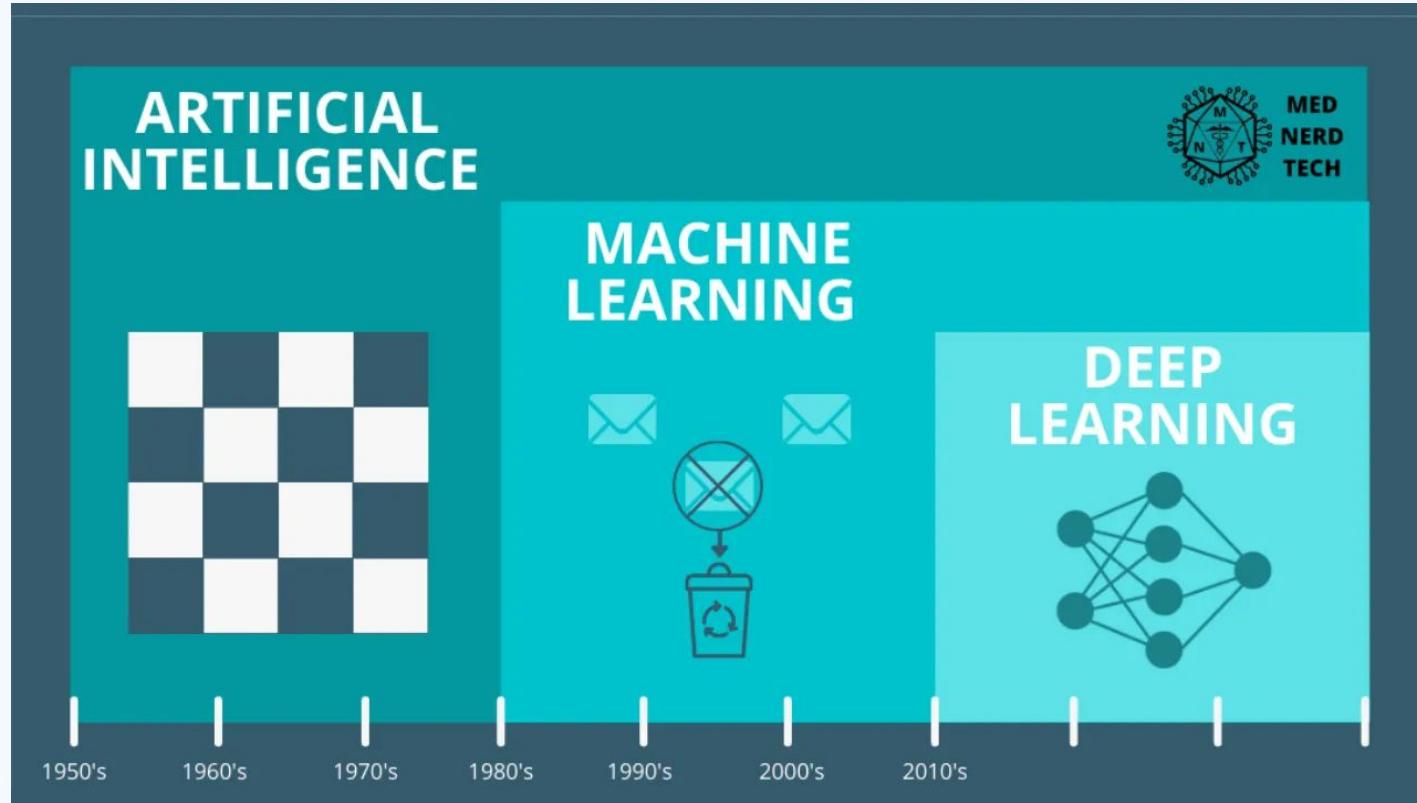
Talvez seja necessário adicionar uma aula a mais, caso o conteúdo atrasse.



02

O que é IA, ML e DL?





Retirado de: <https://medium.com/mednerdtech/ia-x-ml-x-dl-e320867333c1>

O que é Inteligência Artificial?

"A ciência e a engenharia de **fazer máquinas inteligentes**"

– John McCarthy (1955)

"A capacidade de um sistema de **interpretar corretamente dados externos**,
aprender com esses dados e usar esses aprendizados para **alcançar
objetivos e tarefas específicas por meio de adaptação flexível.**"

– Kaplan e Haenlein (2019)



O que é Inteligência Artificial?

Inteligência artificial restrita (Narrow AI):

Refere-se à capacidade de um sistema de computador realizar uma tarefa de forma mais precisa e eficiente do que um humano, **dentro de um escopo bem definido**.

Inteligência artificial geral (AGI):

Em teoria, capaz de resolver **problemas profundamente complexos**, aplicar **julgamento em situações incertas** e incorporar conhecimentos anteriores em seu raciocínio atual, **de criatividade e imaginação equivalentes às dos humanos** e poderia assumir uma gama muito mais ampla de tarefas do que a IA restrita.

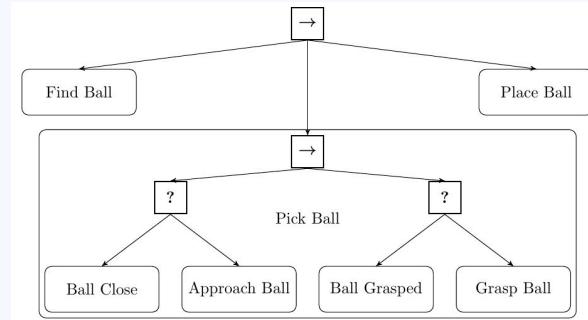
É o tipo de IA que você vê em filmes onde robôs têm pensamentos conscientes e agem com base em seus próprios motivos.

Fonte:

<https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-artificial-intelligence#self-driving-cars>

IA antes de Machine Learning

```
Welcome to      EEEEEE LL   IIII ZZZZZZ AAAAAA  
EE   LL   III ZZZZ AA  AA  
EEEEE LL   II  ZZ  AAAAAAA  
EE   LL   II  ZZ  AA  AA  
EEEEE LLLLLL IIII ZZZZZZ AA  AA  
  
Eliza is a mock Rogerian psychotherapist.  
The original program was described by Joseph Weizenbaum in 1966.  
This implementation by Norbert Landsteiner 2005.  
  
ELIZA: Is something troubling you ?  
YOU: Men are all alike.  
ELIZA: What is the connection, do you suppose ?  
YOU: They're always bugging us about something or other.  
ELIZA: Can you think of a specific example ?  
YOU: Well, my boyfriend made me come here.  
ELIZA: Is it important to you that your boyfriend made you come here ?  
YOU: He says I'm depressed much of the time.  
ELIZA: I'm sorry to hear that you are depressed.  
YOU: It's true. I am unhappy.  
ELIZA: Can you explain what made you unhappy ?  
YOU:
```



Agentes inteligentes baseados em regras:

- Busca de estados
- Min-max
- Behavior Tree
- ...

Exemplos:

- Xadrez, damas, ...
- NPCs em outros tipos de jogos
- Chatbot rudimentares

Retirado de:

<https://en.wikipedia.org/wiki/ELIZA>, <https://www.regencychess.com/blog/artificial-intelligence-and-chess-an-evolving-landscape/>, <https://engineering.nordeus.com/learning-of-behavior-trees-for-autonomous-agents/>

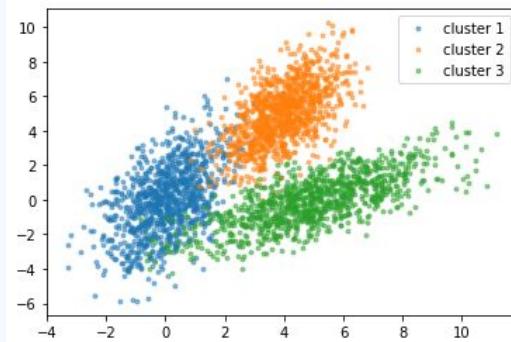
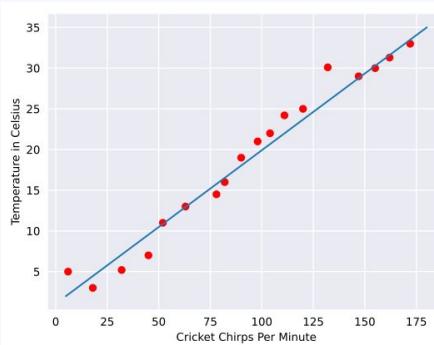
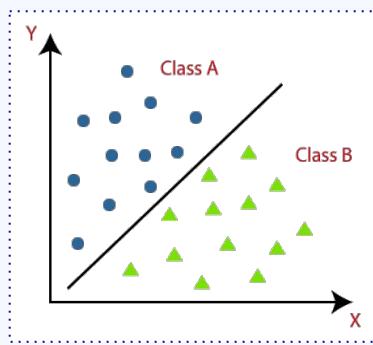
O que é Machine Learning?

“Um programa de computador **aprende a partir de uma experiência E** com respeito a uma **classe de tarefas T** e uma **métrica de performance P**, se a **sua performance em tarefas de T, quando medida por P, melhora com a experiência E.**”

– Tom Mitchell (1997)



O que é Machine Learning?



**Modelos que se
ajustam aos dados:**

- Regressão linear
- Regressão logística
- Algoritmos de Clusterização
- ...

Exemplos:

- Classificadores simples
- Sistemas preditivos
- Sistemas de recomendação
- ...

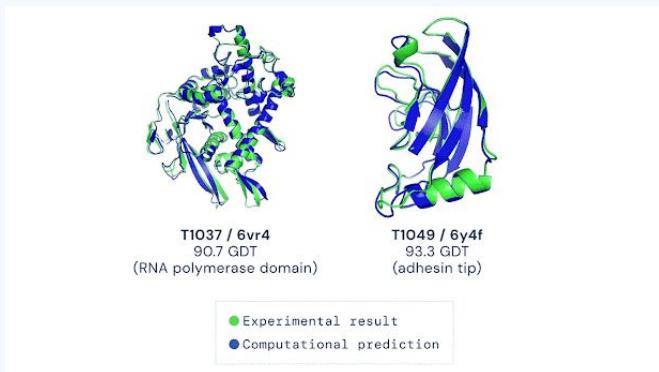
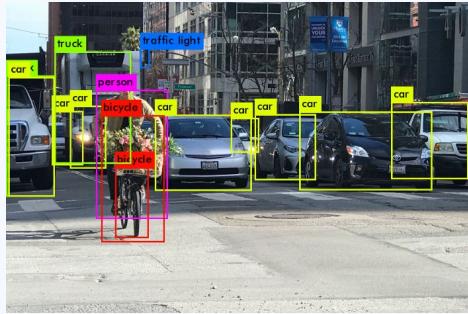
O que é Deep Learning?

“A aprendizagem profunda é uma forma de aprendizagem de máquina que permite aos computadores aprender com a experiência e **compreender o mundo em termos de uma hierarquia de conceitos**. Como o computador adquire conhecimento a partir da experiência, **não é necessário que um operador de computador humano especifique formalmente todo o conhecimento que o computador precisa**. A hierarquia de conceitos permite que o computador **aprenda conceitos complicados construindo-os a partir de conceitos mais simples.**”

– Goodfellow et al. (2016)



O que é Deep Learning?



Retirado de:
<https://exame.com/inteligencia-artificial/o-que-e-chatgpt-como-usar-a-ia-em-portugues-no-seu-dia-a-dia/>
<https://medium.com/analytics-vidhya/yolo-explained-5b6f4564f31>,
<https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/>

Sistemas que descobrem, aprendem e combinam padrões aprendidos:

- CNNs
- RNNs e LSTMs
- Transformers
- ...

Exemplos:

- Sistemas de visão computacional
- Chatbots modernos
- IA generativa
- ...

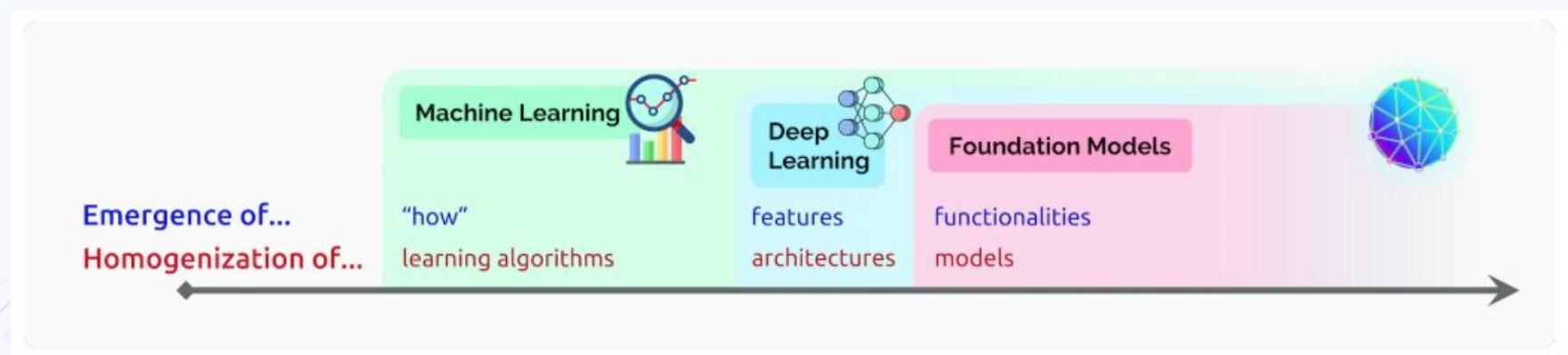
Emergência e homogeneização

Emergência

Surgimento de novos comportamentos.
“Soma das partes não é igual ao todo”.

Homogeneização

Uso de uma mesma técnica em várias áreas.



Emergência e homogeneização

Machine Learning

Emergência

Sistemas passam a ter a capacidade de descobrirem como resolver uma tarefa a partir de dados.

Homogeneização

Diversas aplicações passaram a poder se basear em um algoritmo genérico de aprendizado com base em dados.

Deep Learning

Emergência

Aprendizagem de conceitos abstratos sem a especificação explícita. A escala enorme de modelos trouxe características como *in-context learning*.

Homogeneização

As mesmas arquiteturas de modelos passaram a ser aplicáveis para diversos tipos de aplicações.

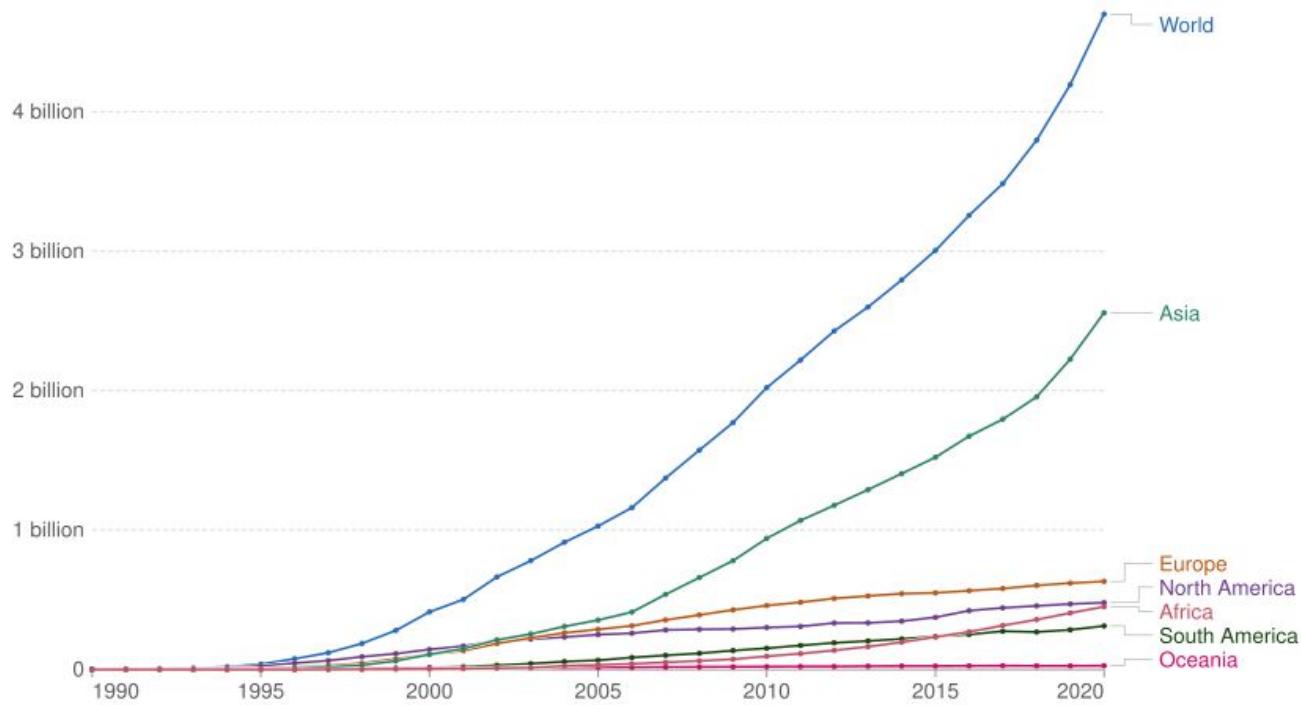
03

A importância dos dados



Number of people using the Internet

Number of people who used the Internet in the last three months.



Data source: OWID based on International Telecommunication Union (via World Bank) and UN (2022)

[OurWorldInData.org/internet](https://ourworldindata.org/internet) | CC BY

Leaders | Regulating the internet giants

The world's most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules



May 6th 2017

Share

ANTONIO GARCÍA MARTÍNEZ

IDEAS FEB 26, 2019 7:08 AM

No, Data Is Not the New Oil

Proposals to "pay" users for the value of their data don't reflect how internet giants like Facebook and Google really operate.



David Parkins

May 6th 2017

Share

A NEW commodity spawns a lucrative, fast-growing industry, prompting antitrust regulators to step in to restrain those who control its flow. A century ago, the resource in question was oil. Now similar concerns are being raised by the giants that deal in data, the oil of the digital era. These titans—Alphabet (Google's parent

ImageNET

14.197.122 de imagens



Retirado de:

<https://rubikscode.net/2021/07/19/top-23-best-public-datasets-for-practicing-machine-learning/>

MegaFace (V2)

672 mil pessoas e 4 milhões de fotos

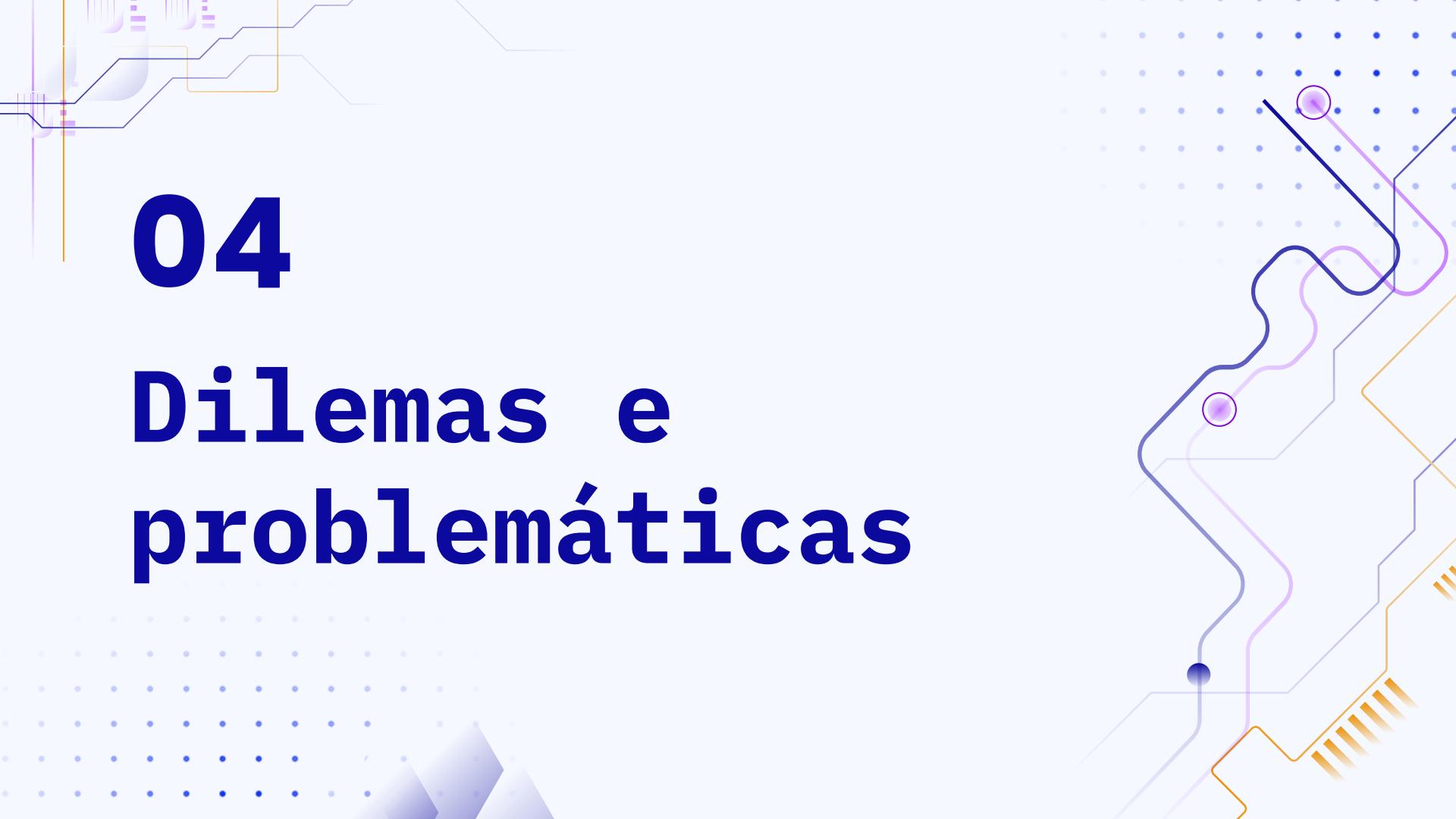


Retirado de:

<https://www.biometricupdate.com/201910/megaface-facial-recognition-dataset-origin-raises-privacy-and-liability-concerns>

04

Dilemas e problemáticas

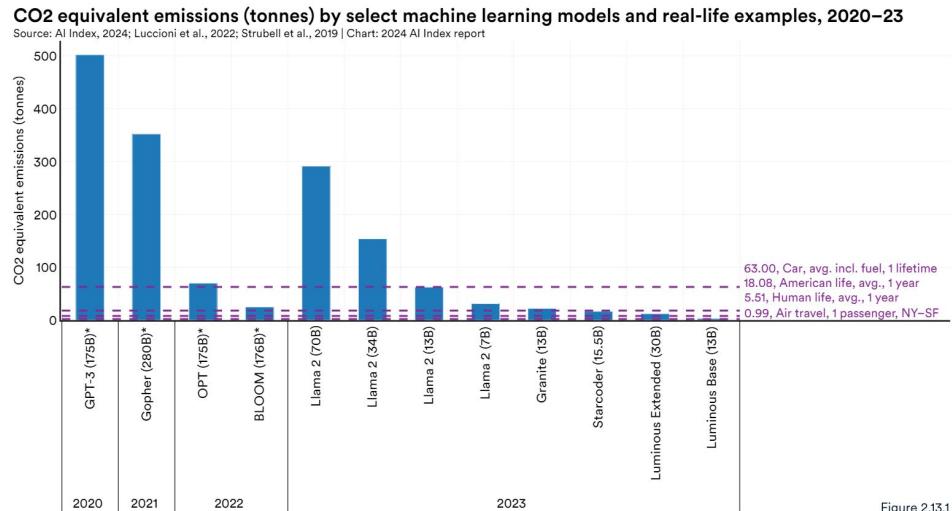


Problemáticas

Custo energético e implicações ambientais.

Usos indevidos e indiscriminados.

Vieses sociais, culturais e políticos.



Retirado de:
<https://aiindex.stanford.edu/report/>

Figure 2.13.1

Problemáticas

Custo energético e implicações ambientais.

Usos indevidos e indiscriminados.

Vieses sociais, culturais e políticos.



Retirado de:
<https://noticias.uol.com.br/cotidiano/ultimas-noticias/2024/04/28/reconhecimento-facial-erros-falta-de-transparencia.htm>

Problemáticas

Custo energético e implicações ambientais.

Usos indevidos e indiscriminados.

Vieses sociais, culturais e políticos.



renatasouzario • Seguir
Rio de Janeiro, Rio de Janeiro

renatasouzario • Racismo algorítmico!

Ao criar uma arte inspirada nos pôsteres da Disney, me deparei com uma imagem gerada a partir de Inteligência Artificial que me retratava como uma mulher negra com uma arma na mão. A descrição pedida era de uma mulher negra, de cabelos afro, com roupas de estampa africana num cenário de favela. E essa foi a imagem gerada. Não pode uma mulher negra, crie a favela, estar num espaço que não da violência? O que leva essa "desinteligência artificial" a associar o meu corpo, a minha identidade, com uma arma?

Em um mundo forjado pela branquitude, que tem o domínio

4.613 curtidas
HÁ 2 HORAS

Entrar para curtir ou comentar.

Retirado de:

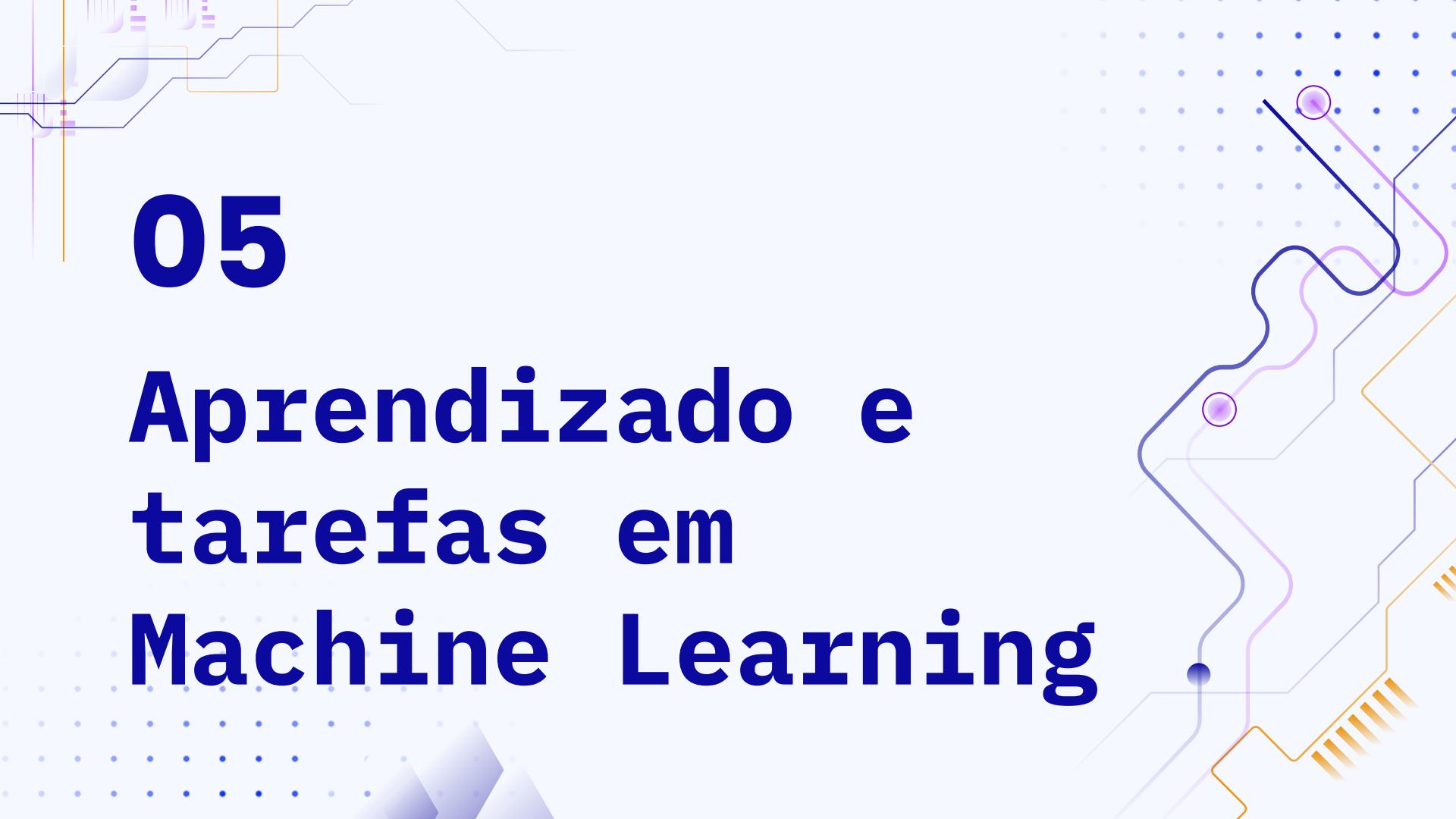
<https://g1.globo.com/ri/rio-de-janeiro/noticia/2023/11/09/apos-denuncia-de-racismo-em-ferramenta-de-inteligencia-artificial-deputada-foi-chamada-de-macaca-em-redes-sociais.ghtml>



AI, Ain't I A Woman? de Joy Buolamwini

05

Aprendizado e tarefas em Machine Learning

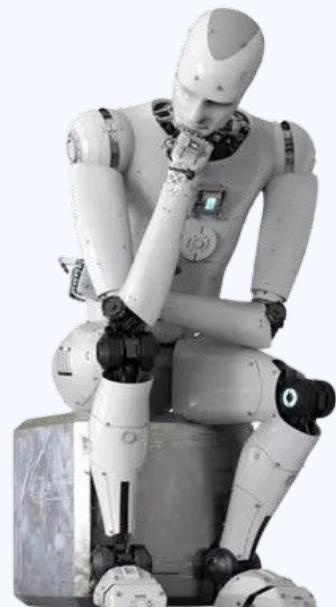


O aprendizado de máquina

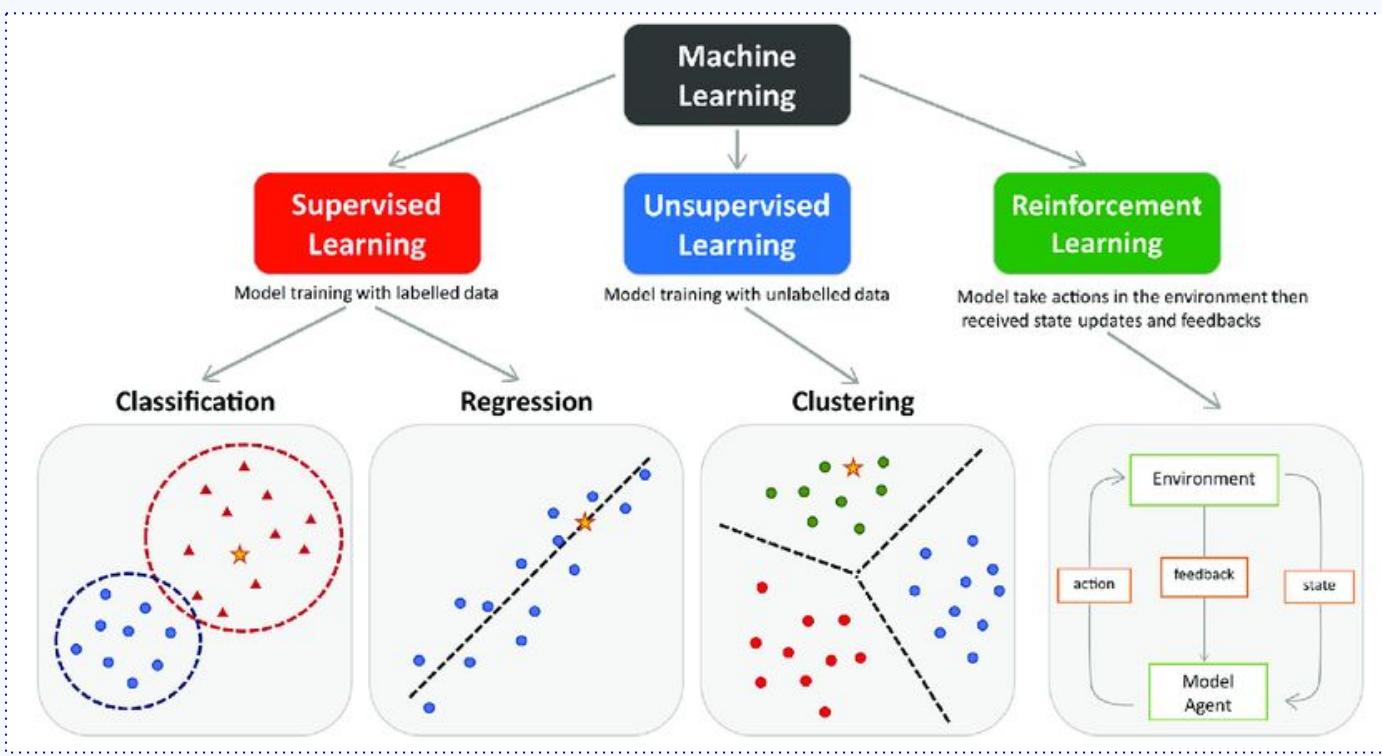
Programação tradicional:



Aprendizado de máquina:



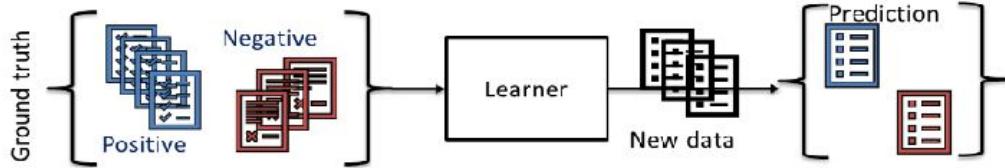
Paradigmas de aprendizado



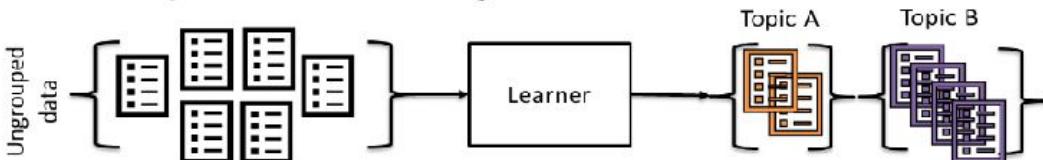
Retirado de: https://www.researchgate.net/figure/The-main-types-of-machine-learning-Main-approaches-include-classification-and_fig1_354960266

Paradigmas de aprendizado

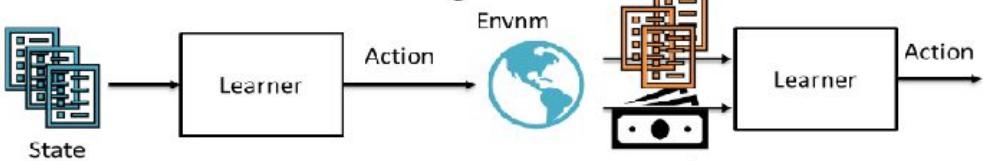
❖ Supervised Learning



❖ Unsupervised Learning



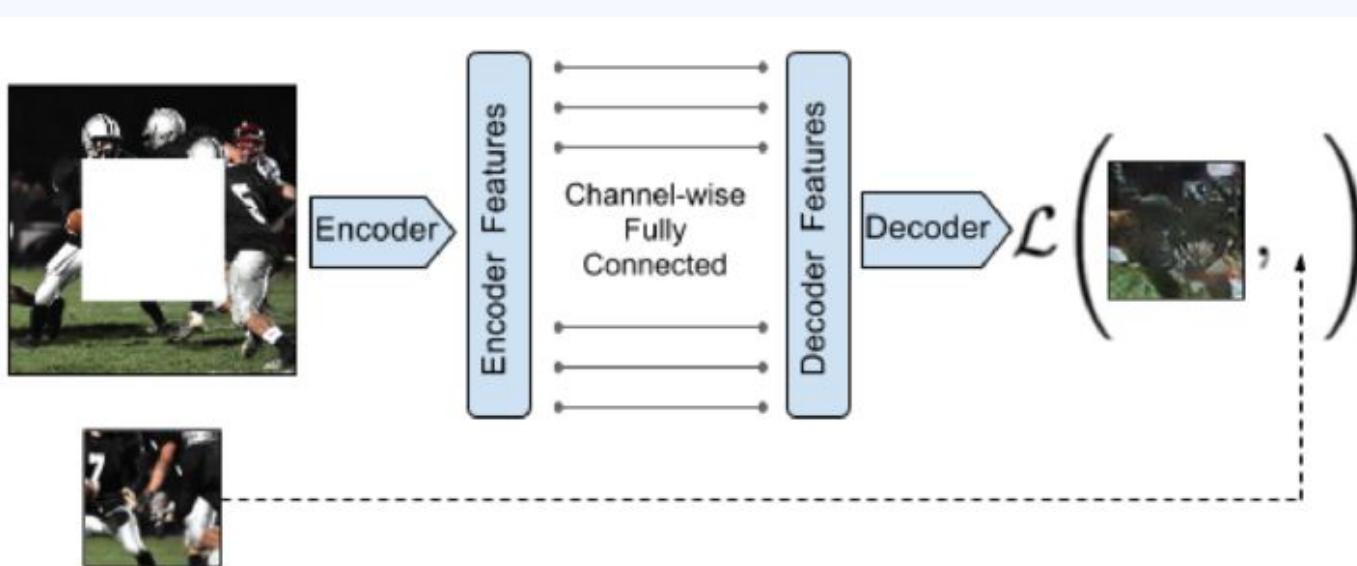
❖ Reinforcement Learning



Paradigmas de aprendizado

Importante mencionar um paradigma, usado para treinamento de sistemas de Deep Learning, que surgiu última década...

Aprendizado Auto-supervisionado



Paradigmas de aprendizado

Aprendizado supervisionado:



O objetivo é “**predizer**” uma saída esperada/alvo.



Saída esperada /alvo é **conhecida**.

Aprendizado Não-supervisionado:



O objetivo é **encontrar padrões** na estrutura dos dados.



Não há uma saída esperada /alvo **conhecida**.

Aprendizado Auto-supervisionado:

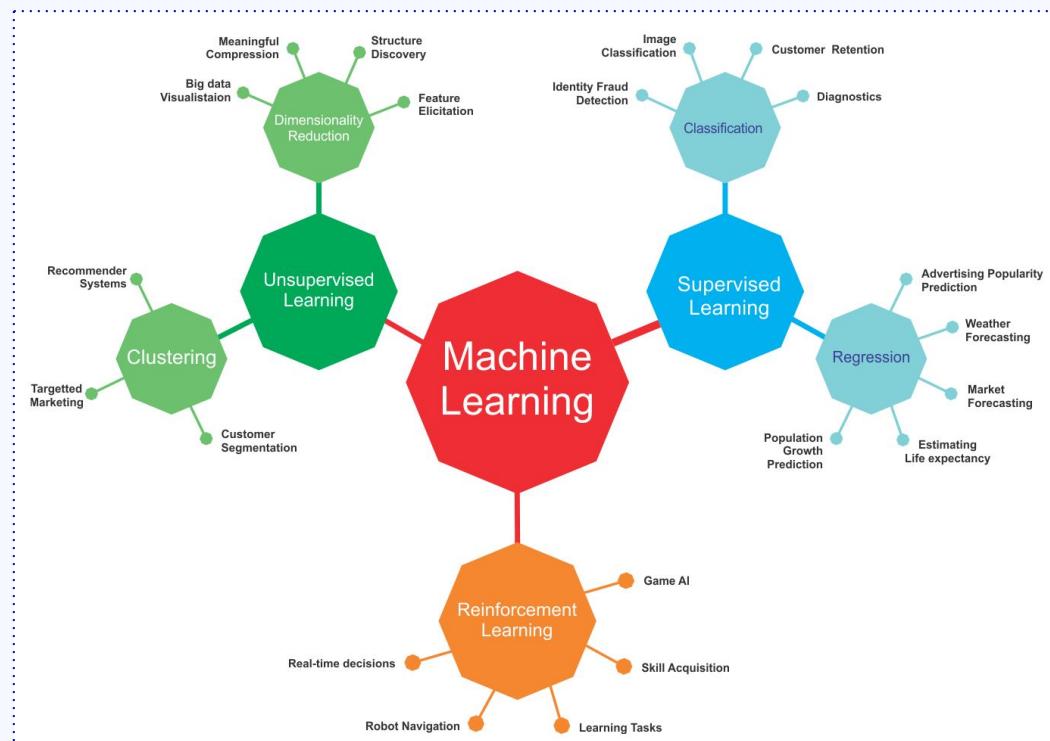


O objetivo é extrair e **capturar os padrões** nos dados.



A saída esperada /alvo **faz parte dos dados de entrada**.

Tarefas (ou “problemas”)



Retirado de: <https://subscription.packtpub.com/book/data/9781789345070/1/ch01lv1sec04/ml-tasks>

06

Exercícios de revisão



Revisão e atividade prática

Ambiente de desenvolvimento: Google Colab.

Atividade de revisão:

- Python;
- Manipulação de dados;
- Análise de dados;
- Operação com matrizes.



Referências

Parte do material foi inspirado nos slides da Prof^a. Sandra Avila e do Prof. Anderson Rocha.

Livros utilizados:

- *Deep Learning* (2016)
- *An Introduction to Statistical Learning* (2023)
- *Pattern Recognition and Machine Learning* (2006)
- *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2017)

Obrigado pessoal!

Até próxima aula :)



Iris Data Science UNICAMP



@irisdatascienceunicamp

