



MACHINE LEARNING APLICADO À ANÁLISE DE DADOS

AULA 1 – 15/03/2021

EQUIPE TÉCNICA

- Profa. Dra. Hieda Adriana Nascimento Silva (Coordenadora)
- Prof. Dr. Marcos César da Rocha Seruffo (Subcoordenador)
- Profa. Dra. Yomara Pinheiro Pires (Subcoordenadora)
- Profa. Dra. Jakelyne Machado Lima Silva (Subcoordenadora e participante externa)
- Alana Miranda Medeiros (Bolsista)
- Lucas Sousa dos Santos (Monitor e voluntário)
- Suze Blenda de Paula Lima (Voluntária)
- Vitor Lima Costa Neto (Voluntário)

OBJETIVO GERAL

- Oferecer cursos sobre Machine Learning à comunidade em geral, englobando profissionais de diversas áreas, técnicos, alunos e professores, capacitando-lhes na área de aprendizagem de máquina, de forma que possam aplicar esses estudos em seus diversos projetos, disseminando assim a inteligência artificial para comunidade interna e externa à UFPA e tornando os projetos mais otimizados e alinhados com as tendências científicas.

OBJETIVOS ESPECÍFICOS

- Disseminar o conhecimento por meio do estudo de Machine Learning aplicada à análise de dados;
- Integrar os profissionais de diferentes áreas e de outras instituições, a partir de atividades extensionistas como formação e treinamentos na área de Machine Learning;
- Despertar interesse nos docentes e discentes envolvidos para que entendam a importância do uso da inteligência artificial em ambientes interdisciplinares;
- Gerar documentos (artigos e/ou relatórios) com os resultados obtidos dos cursos, para que o estudo da Inteligência Artificial e Machine Learning possam efetivamente impactar a sociedade, contribuindo para a tomada de decisão;

COMPETÊNCIAS E HABILIDADES

Competências:

Aplicar conhecimentos básicos das principais técnicas de Machine Learning em diversas áreas do conhecimento.

Habilidades:

Analisar os diversos problemas e definir a melhor técnica de ML a ser utilizada para sua implementação

EMENTA

- Introdução aos conceitos de ML
- Google Colab e fundamentos Python
- Análise de dados com Numpy e Pandas
- Visualização de dados com Matplotlib e Seaborn
- Introdução a scikit-learn
- Regressão linear e regressão logística
- Árvore de decisão
- Clusterização com k-means
- Redes Neurais Artificiais
- Projeto Final

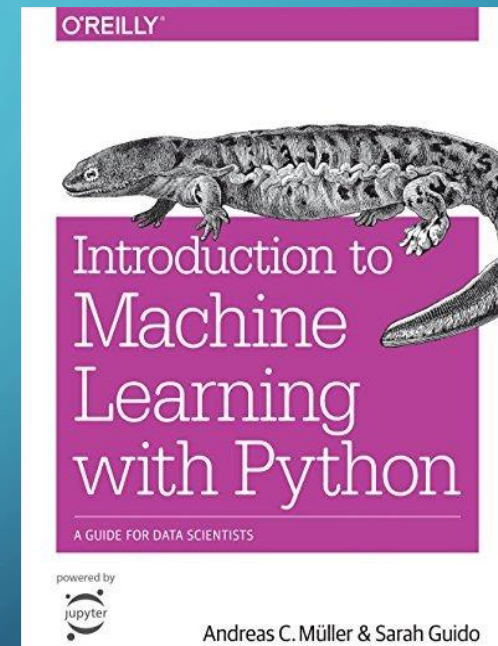
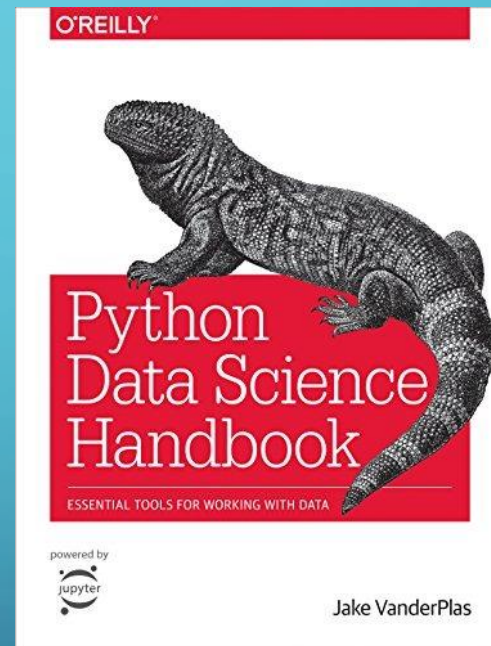
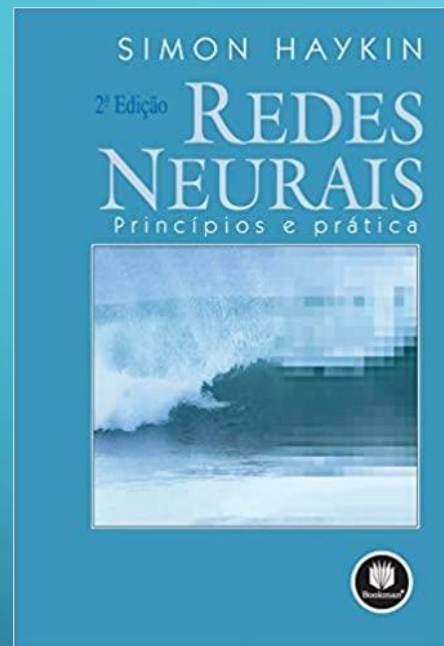
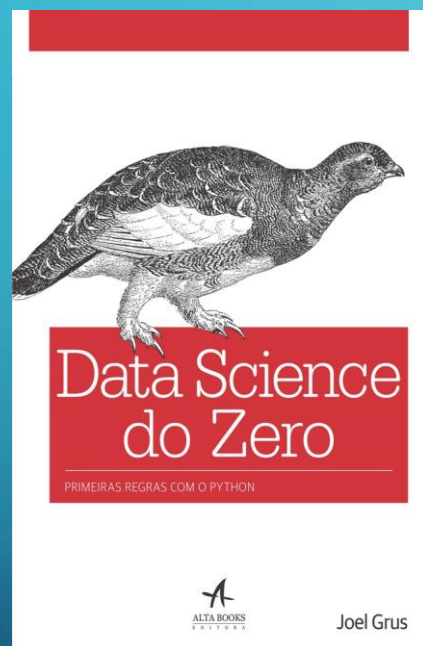
AVISOS

- Dias e horários das aulas
- Formato das aulas: Teoria + prática no Google Colab
- Atividades
- Certificado: 75% de frequência
- Canal de Comunicação: Telegram
- Questionário avaliativo

DICAS PARA O CURSO

- Estude Python
- Anote o que achar importante nas aulas
- Pratique muito
- Use o Github

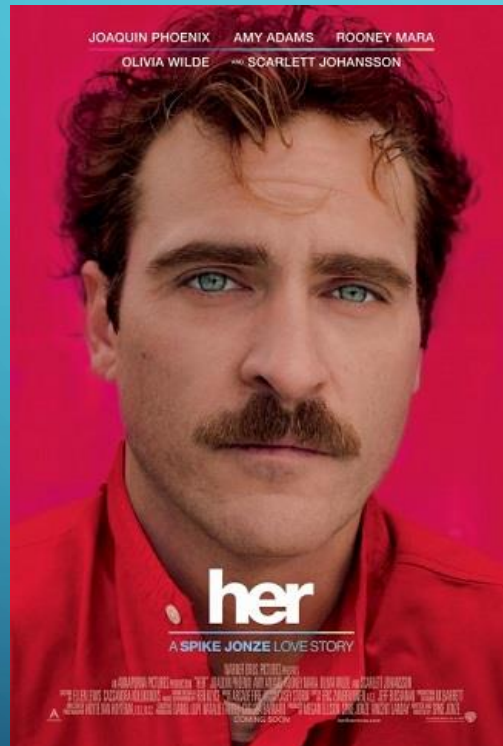
INDICAÇÕES DE LIVROS



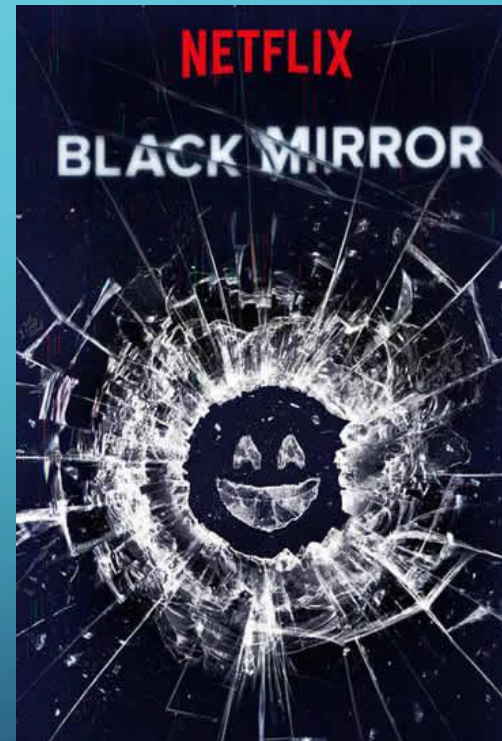
INDICAÇÕES DE CURSOS ONLINE

- PYTHON FUNDAMENTOS PARA ANÁLISE DE DADOS (DATA SCIENCE ACADEMY)
- PYTHON FOR DATA SCIENCE AND MACHINE LEARNING BOOTCAMP (UDEMY)
- MACHINE LEARNING - STANFORD (COURSERA)

FILMES E SÉRIES SOBRE INTELIGÊNCIA ARTIFICIAL/MACHINE LEARNING



(2013)



(2011)

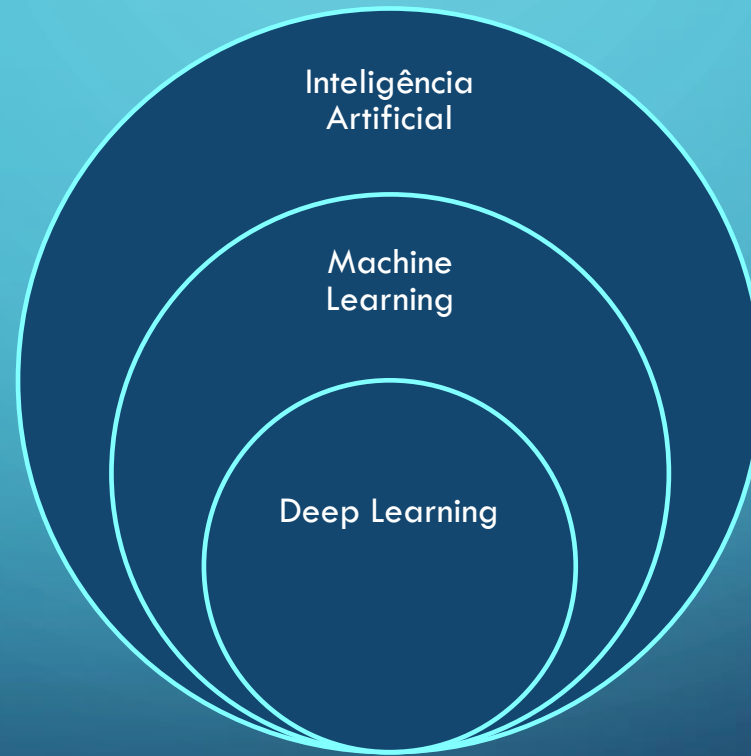
O QUE É MACHINE LEARNING (APRENDIZADO DE MÁQUINA)?

“É o campo de estudo que dá aos computadores a capacidade de aprender sem serem explicitamente programados.”

Arthur Samuel – cientista da computação e pioneiro nos campos de IA e Machine Learning



MACHINE LEARNING X INTELIGÊNCIA ARTIFICIAL X DEEP LEARNING



O QUE É UM DATASET?

- É um arquivo que contém dados sobre um determinado assunto. São organizados em linhas e colunas e utilizados para análise de dados e estatística.

Formatos: xls (planilha do excel), csv, txt, json e xml.

ONDE ENCONTRAR DATASETS?

Para começar:

- **Kaggle** (<https://www.kaggle.com/>)
- **UCI Machine Learning Repository** (<https://archive.ics.uci.edu/ml/index.php>)

Dados do Brasil:

- **Portal Brasileiro de Dados Abertos** (<https://dados.gov.br/>)
- **Portal da Transparência** (<http://www.portaldatransparencia.gov.br/>)

ONDE ENCONTRAR DATASETS?

Dados da saúde:

- **Global Health Observatory** (<https://www.who.int/data/gho>)
- **CORD-19** (<https://www.semanticscholar.org/cord19>)

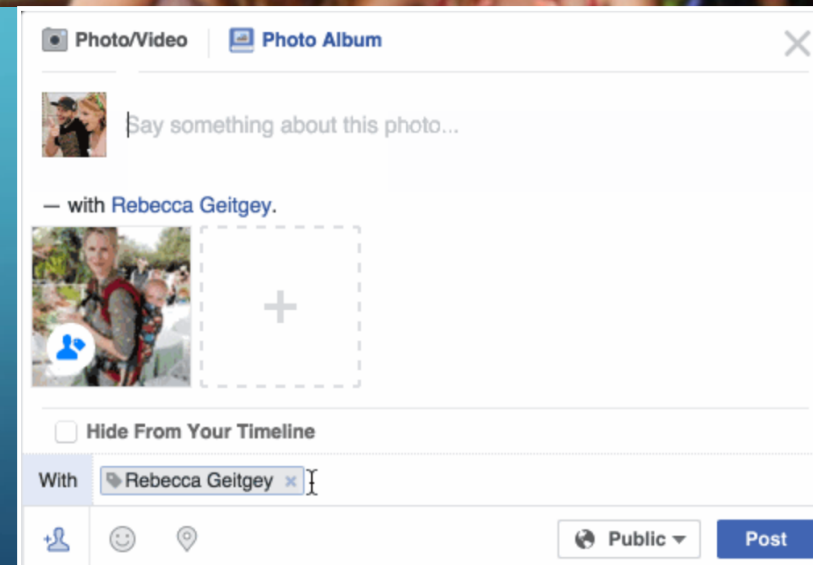
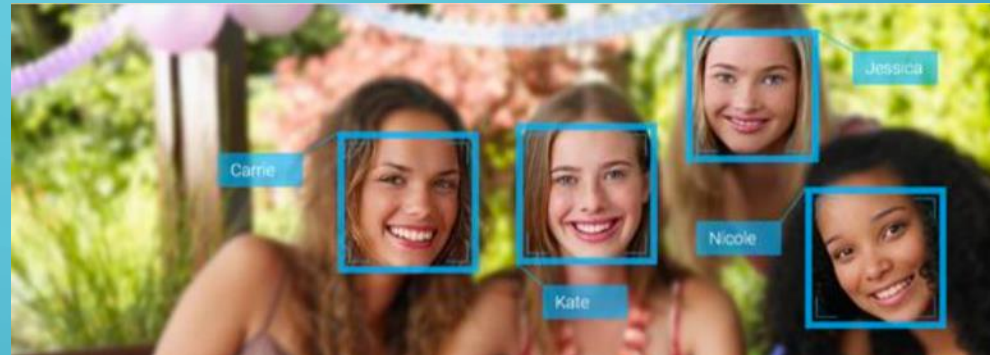
Outros dados:

- **Quandl** (<https://www.quandl.com/search>)
- **FiveThirtyEight** (<https://data.fivethirtyeight.com/>)

APLICAÇÕES COM MACHINE LEARNING

RECONHECIMENTO FACIAL

Desenvolvido para identificar, detectar e rastrear uma pessoa por meio de imagem ou vídeo.



SISTEMAS DE RECOMENDAÇÃO

São capazes de filtrar o conteúdo que será entregue aos usuários com base em seus gostos e preferências.

[Help](#) | [Close window](#)

Recommended for You



Inside Apple: How America's Most Admired--and Secretive--Company Really Works
Our Price: **\$9.99**
Used & new from **\$9.99**

[See all buying options](#)


Rate this item

☒ ★★★★★

☐ I own it

☐ Not interested

Because you purchased...



The Toyota Way : 14 Management Principles from the World's Greatest Manufacturer
(Kindle Edition)

☒ ★★★★★

☐ This was a gift

☐ Don't use for recommendations

 **NETFLIX**



O Gambito da Rainha

98% relevante 2020 16 Minissérie 5.1

T1:E1 Aberturas

Com nove anos, Beth é enviada a um orfanato e lá desenvolve dois traços: o talento incrível para o xadrez e a dependência do tranquilizante dado às crianças.

Elenco:Anya Taylor-Joy, Bill Camp, Marielle Heller

Criação:Scott Frank, Allan Scott

[+](#) [👍](#)

Minha lista Classificar

MINISSÉRIE

Populares na Netflix

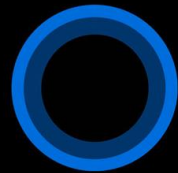
Assistir novamente

BROOKLYN NINE-NINE

THE DARK

PROCESSAMENTO DE LINGUAGEM NATURAL

Técnica computacional usada para ajudar dispositivos tecnológicos a entender, interpretar e manipular a linguagem humana.



Hi. I'm Cortana.



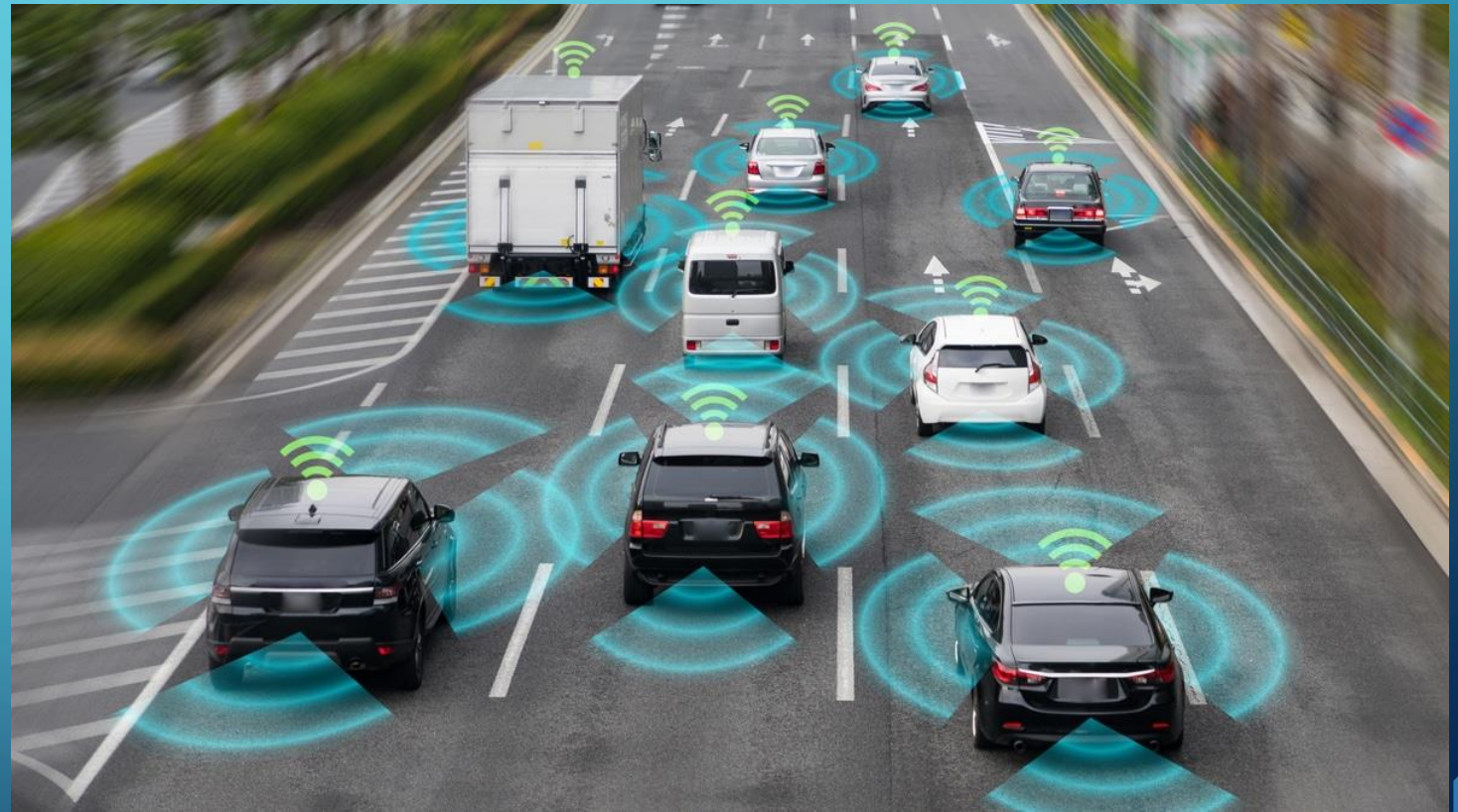
DIAGNÓSTICO DE CUIDADOS DE SAÚDE

- Análise de imagens de exames radiológicos e detecção de anomalias
- Detecção de doenças raras
- Associação de sintomas
- Redução da taxa de erro no diagnóstico de Câncer



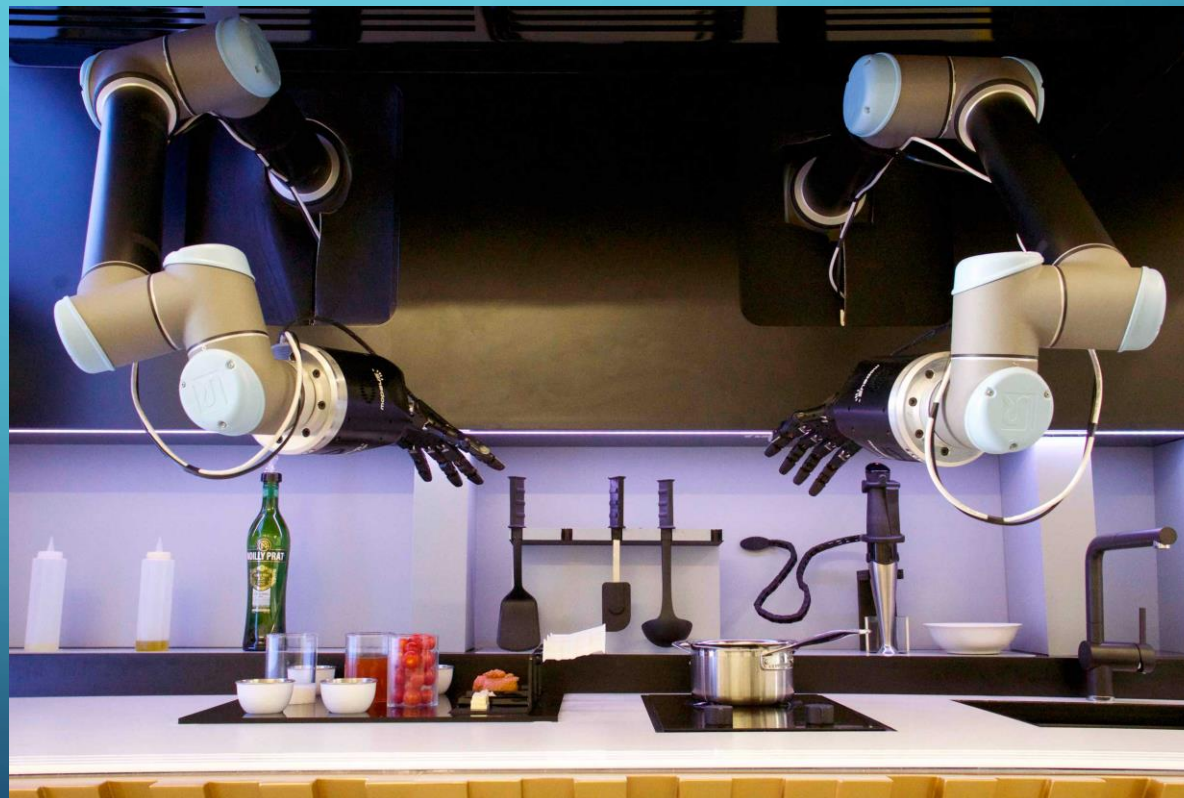
VEÍCULOS AUTÔNOMOS

A ideia dessa nova tecnologia é fazer com que os veículos possam visualizar o mundo real e identificar objetos no espaço em que estão inseridos e classificá-los em dois tipos: estáticos e dinâmicos.



ROBÔS

Robôs com habilidades de aprendizado de máquina podem melhorar suas capacidades ao longo do tempo.



APRENDIZAGEM SUPERVISIONADA X NÃO SUPERVISIONADA

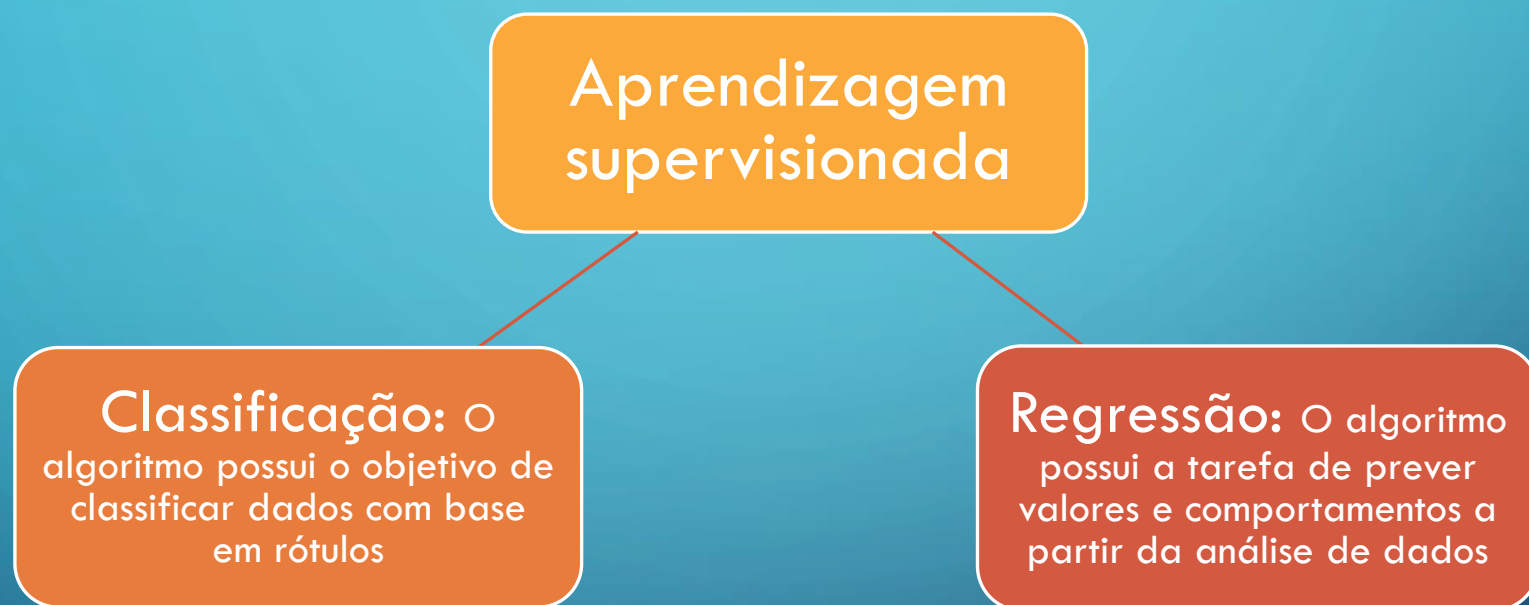
APRENDIZAGEM SUPERVISIONADA

- Previsão de valores ou classes
- Dados de treino precisam conter valores de entrada e de saída, pra que o modelo aprenda como, a partir de novos dados de entrada, gerar a saída correta

APRENDIZAGEM NÃO SUPERVISIONADA

- Identificação de grupos (clusters) de dados
- Os dados de treino contem apenas os dados de entrada

APRENDIZAGEM SUPERVISIONADA



APRENDIZAGEM NÃO SUPERVISIONADA

Aprendizagem não supervisionada

Agrupamento(clustering):

O algoritmo possui a tarefa de separar dados em grupos, segmentando por características similares

REFERÊNCIAS BIBLIOGRAFICAS

- [17 Casos de Uso de Machine Learning - Data Science Academy](http://datascienceacademy.com.br/blog/17-casos-de-uso-de-machine-learning/) (<http://datascienceacademy.com.br/blog/17-casos-de-uso-de-machine-learning/>)
- [Anaconda Python \(datascienceacademy.com.br\)](https://www.datascienceacademy.com.br/path-player?courseid=python-fundamentos&unit=5722d08747d7dd30ac8b457bUnit) (<https://www.datascienceacademy.com.br/path-player?courseid=python-fundamentos&unit=5722d08747d7dd30ac8b457bUnit>)
- [Início - Deep Learning Book](http://deeplearningbook.com.br/deep-learning-a-tempestade-perfeita/) (<http://deeplearningbook.com.br/deep-learning-a-tempestade-perfeita/>)
- [Os 7 melhores sites para encontrar datasets para projetos de Data Science | by Paulo Vasconcellos](https://paulovasconcellos.com.br/os-7-melhores-sites-para-encontrar-datasets-para-projetos-de-data-science-8a53c3b48329) (<https://paulovasconcellos.com.br/os-7-melhores-sites-para-encontrar-datasets-para-projetos-de-data-science-8a53c3b48329>)