

Machine Learning

Aprendizado de Máquina

Vinícius Quintão Carneiro Vinicius.carneiro@ufla.br



Machine Learning Aprendizado de Máquina

Conjunto de técnicas usadas com o intuito de programar computadores para aprender e reconhecer padrões automaticamente, sem serem explicitamente programados (modelados).



Aprendizado e Generalização

| OBS | Altura | Som | Alimento | Animal Doméstico | Produto Comercial |
|-----|--------|--------|---------------|---------------------|----------------------|
| 1 | 70 cm | Latido | Ração | Sim | Não |
| 2 | 1,50 m | mugido | Capim | sim | Leite |
| 3 | 30 cm | Chia | Ovos e Frutas | sim | Não |

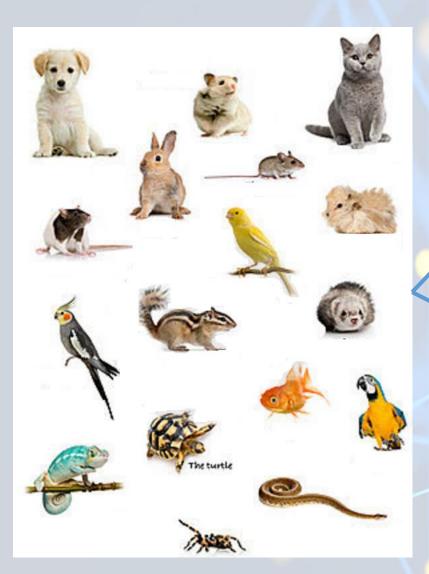
Aprendizado e Generalização

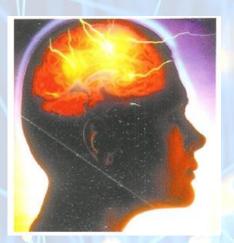
















Que animal é este ???



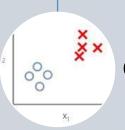
Machine Learning (Aprendizado de Máquina)



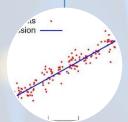
Supervisionado



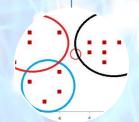
Não Sup<mark>ervisio</mark>nado



Classificação



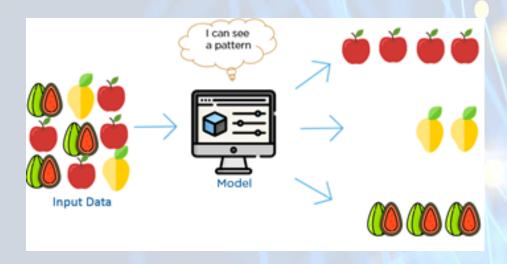
Regressão

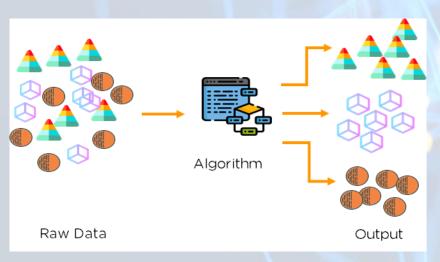


Agrupamento

Redes Neurais Artificiais
Redes de Base Radial
Árvores de Decisão (Algorítmos Ensemble)
K-Vizinhos mais próximos
Máquina de vetores de suporte

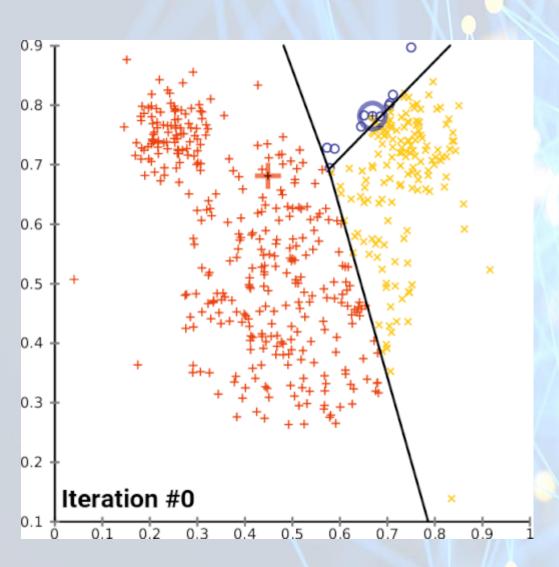
K-means
Fuzzy C-means
Mapas Auto Organizáveis
Mean Shift
Agrupamento Espectral





http://hunterheidenreich.co m/blog/breaking down ml for the average person/

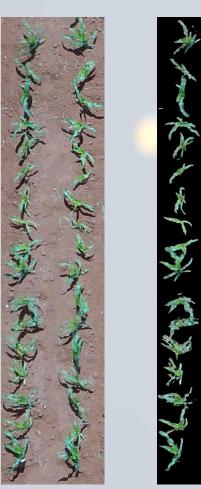
http://shabal.in/visuals/kmeans/2.html





| 42 | 23 | 43 | 54 | 48 | 61 | 255 | 255 | 69 | 60 | 34 | 30 |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----------|-----|----|
| 24 | 0 | 15 | 19 | 8 | 16 | 255 | 255 | 17 | 13 | 0 | 0 |
| 20 | | | | | | 255 | 255 | 3 | | | |
| 38 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | |
| 10 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 49 | 255 | 155 | 218 | 237 | 243 | 244 | 246 | 216 | 169 | 255 | 42 |
| 5 | 255 | | 147 | 157 | 158 | 152 | 156 | | | 255 | 0 |
| 2 | 0 | 70 | 119 | 120 | 117 | 111 | 119 | 102 | 72 200 | 0 | |
| 59 | 255 | 211 | 249 | 234 | 231 | 241 | 253 | 243 | 200 | 255 | |
| 0 | 255 | | 158 | 136 | 128 | 134 | 147 | 145 | | 255 | 0 |
| 0 | 0 | 106 | 129 | 97 | 83 | 88 | 105 | 110 | 88 | 0 | |
| 67 | 255 | 238 | 255 | 230 | 236 | 250 | 255 | 252 | 224 | 255 | |
| 0 | 255 | 135 | 147 | 114 | 115 | 128 | 137 | 140 | 125 | 255 | |
| 0 | 0 | 116 | 118 | 73 | 68 | 79 | 92 | 102 | 94 | 0 | |
| 92 | 255 | 238 | 255 | 245 | 255 | 255 | 250 | 255 | 240 | 255 | |
| 14 | 255 | 128 | 143 | 122 | 130 | 130 | 124 | 142 | 137 | 255 | 8 |
| 12 | 0 | 111 | 114 | 81 | 82 | 79 | 76 | 102 | 104 | 0 | |
| 93 | 255 | 206 | 245 | 249 | 255 | 252 | 244 | 253 | 237 | 255 | |
| 22 | 255 | | 136 | 131 | 135 | 128 | 122 | 140 | 137 | 255 | |
| 20 | 0 | 86 | 107 | 91 | 88 | 77 | 73 | 98 | 105 | 0 | |
| 61 | 255 | 163 | 222 | 236 | 244 | 249 | 255 | 232 | 218 | 255 | |
| 1 | 255 | | 127 | 132 | 136 | 137 | 148 | | 127 | 255 | 10 |
| 0 | 0 | 58 | 99 | 93 | 90 | 89 | 103 | 89 | 98 | 0 | |
| 57 | 255 | 104 | 193 | 252 | 253 | 245 | 250 | 223 | | 255 | |
| 15 | 255 | 41 | 121 | 172 | 168 | 156 | 161 | | | 255 | 16 |
| 16 | 0 | 26 | 96 | 137 | 127 | 114 | 121 | 104 | 45 | 0 | |
| 38 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | |
| 14 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 14 28 | | 19 | 25 | | | | | | 34 | | |
| 18 | 10 | | 0 | 10 | 21 | 17 | | 4 | 0 | 0 | 4 |
| 19 | | | | | | | | | | | |
| | | | | | | | | | | | |

| Pixel | R | G | В |
|-------|-----|-----|-----|
| 1 | 42 | 24 | 20 |
| 2 | 23 | 0 | 0 |
| 3 | 43 | 15 | 0 |
| 4 | 54 | 19 | 0 |
| 5 | 48 | 8 | 0 |
| ••• | ••• | ••• | ••• |
| 132 | 33 | 4 | 0 |





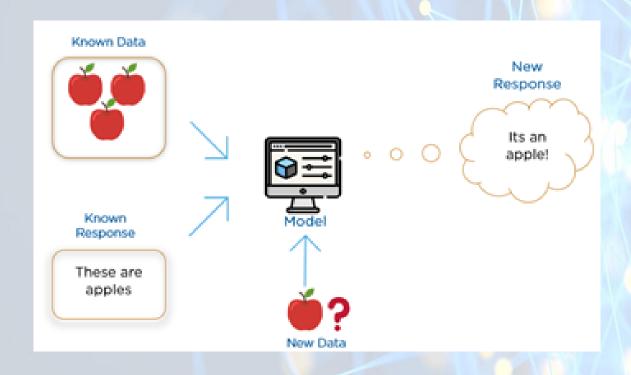


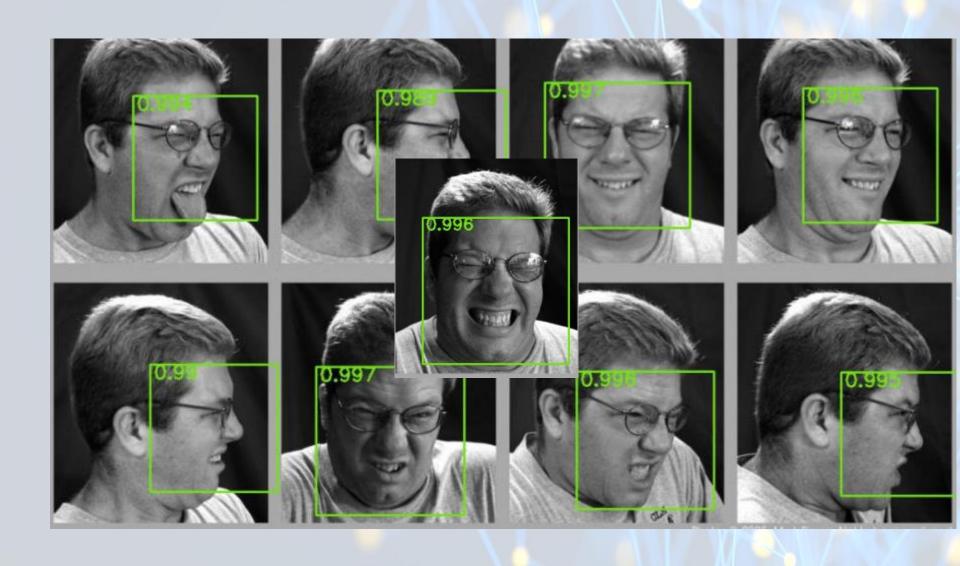


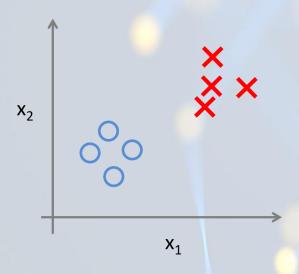


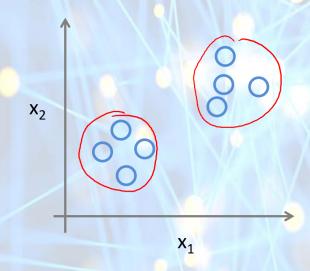
Etapas de desenvolvimento de um sistema ML Aprendizado Não Supervisionado

- 1. Definir o problema
- 2. Escolha das variáveis de entrada
- 3. Escolha do algoritmo de ML > Não supervisionado
- 4. Definição dos parâmetros do algoritmo escolhido: Número de Grupos
- 5. Uso do algoritmo







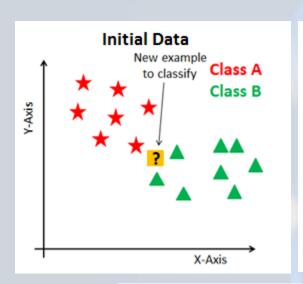


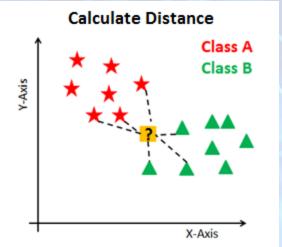
| Observas e | | Saída Real | | |
|--------------|---------------|-------------------|-------------|------------|
| Observação - | Produtividade | Altura de Plantas | Resistência | Salda Real |
| 1 | 1000 | 20 | 1 | 1 |
| | | | | |
| 1000 | 1500 | 40 | 3 | 4 |

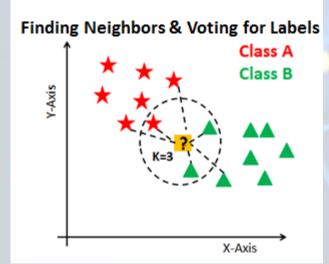
| Observação | Entradas | | | | | |
|------------|---------------|-------------------|-------------|--|--|--|
| | Produtividade | Altura de Plantas | Resistência | | | |
| 1 | 1000 | 20 | 1 | | | |
| | ••• | ••• | | | | |
| 1000 | 1500 | 40 | 3 | | | |

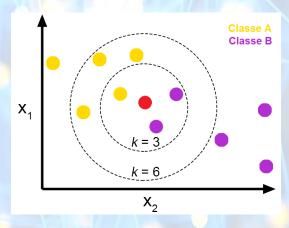
Principais Algoritmos de ML

1. K-Nearest Neighbors (KNN)







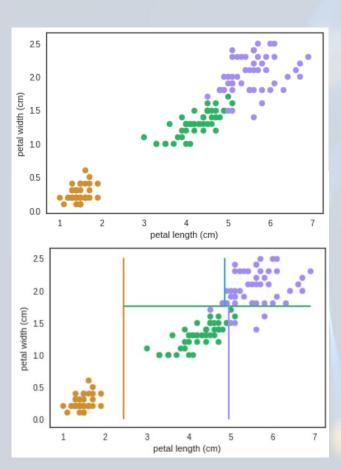


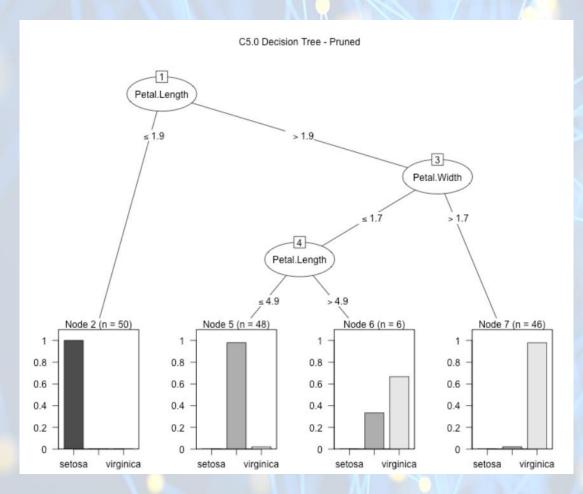
https://inferir.com.br/artigos/algoritimo-knn-para-classificacao/



Principais Algoritmos de ML

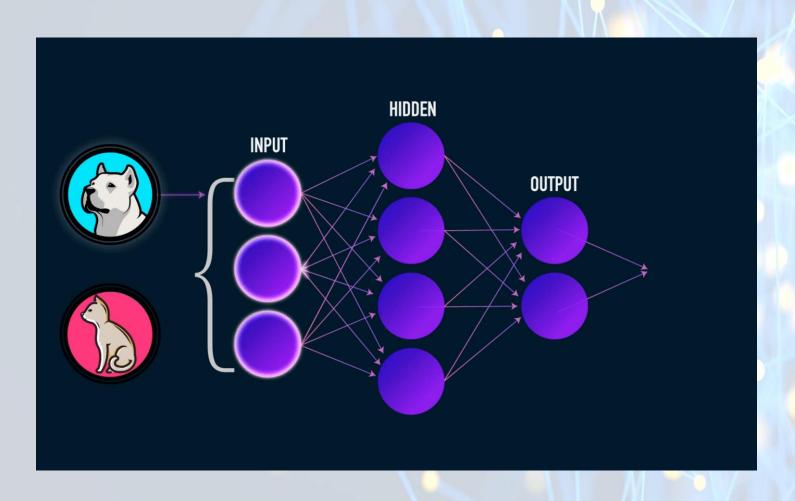
2. Árvores de Decisão



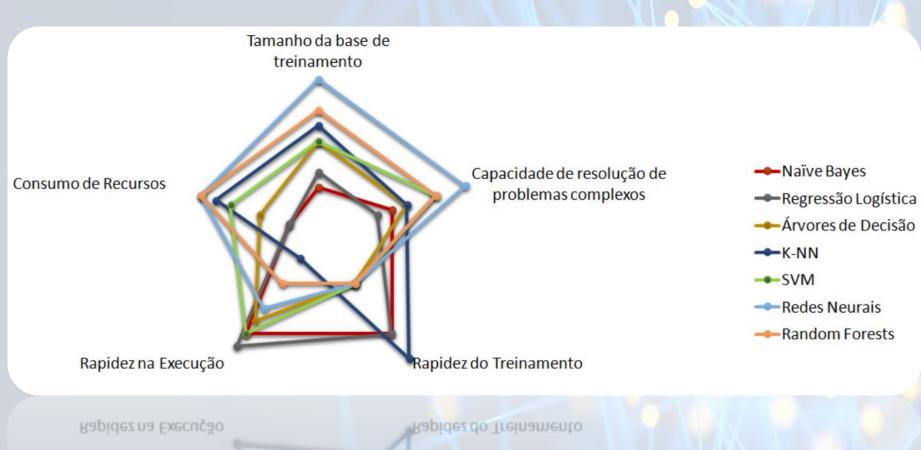


Principais Algoritmos de ML

3. Redes Neurais Artificiais



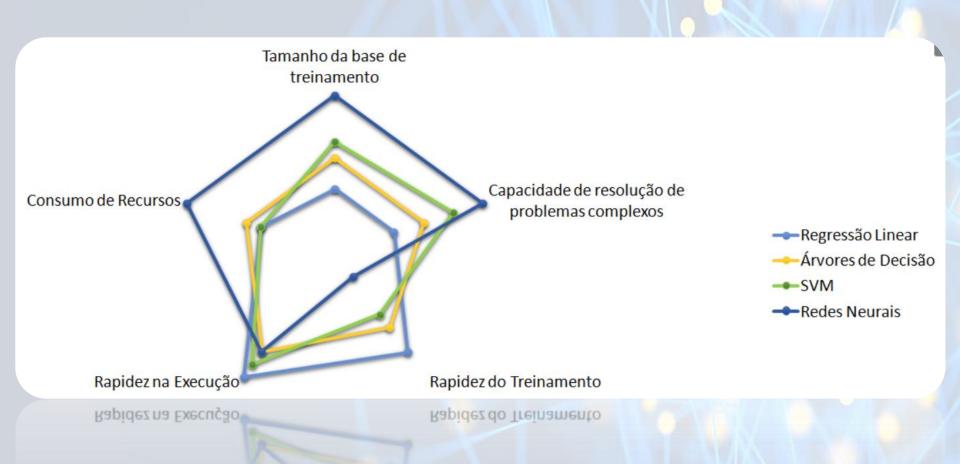
Comparação - Classificação



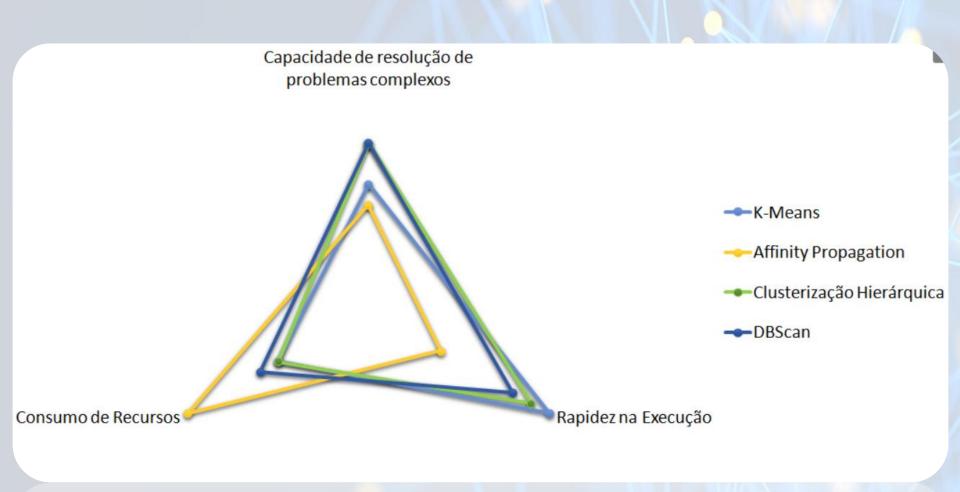
-Random Forests

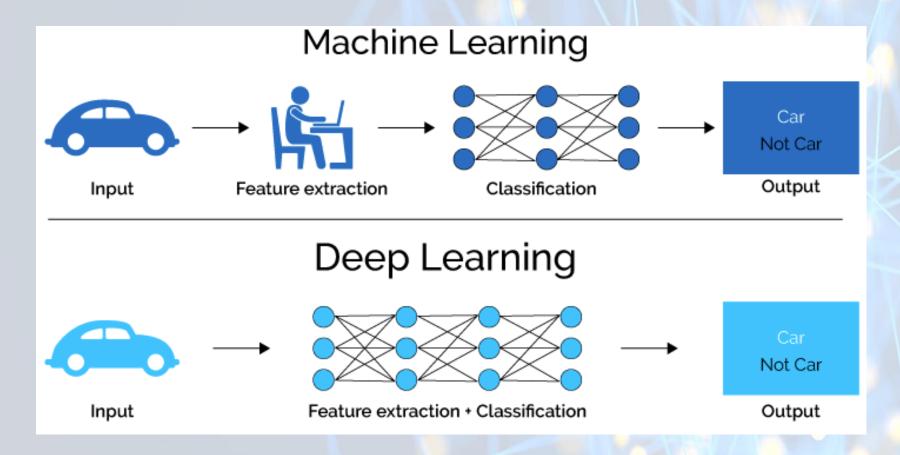
Redes Neurals

Comparação - Regressão

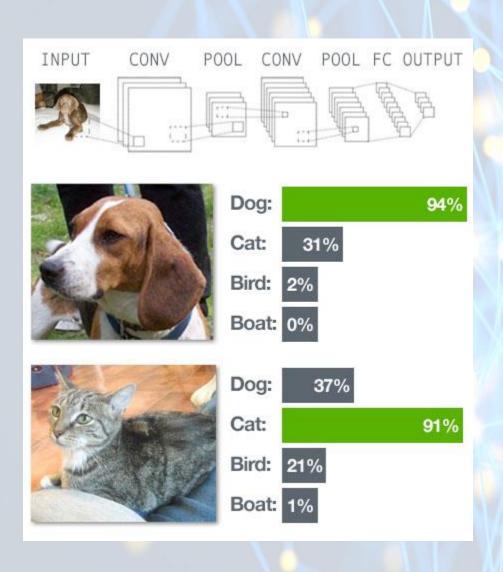


Comparação - Agrupamento





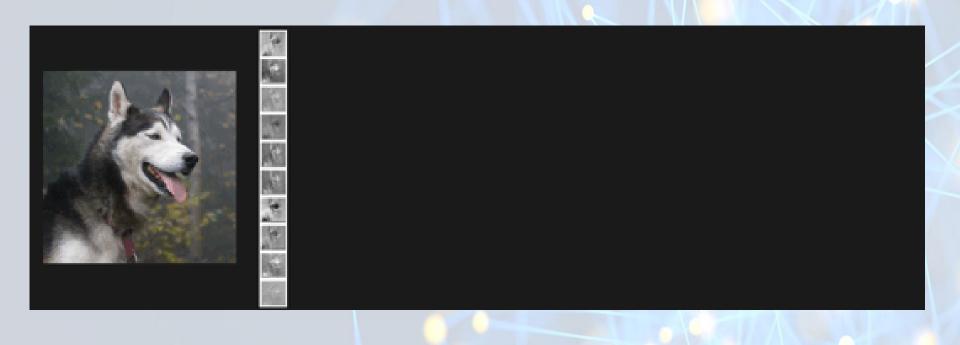
Redes Neurais Convolucionais



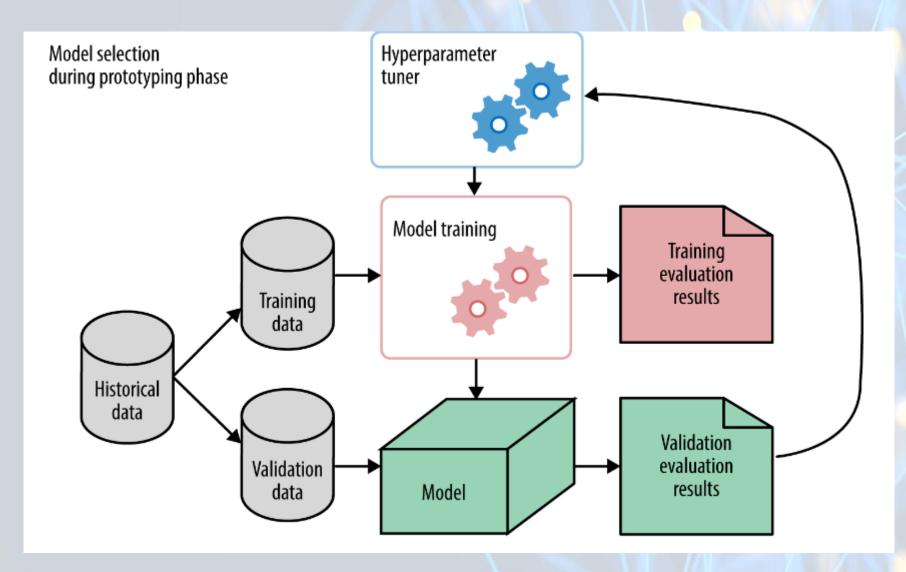
Redes Neurais Convolucionais

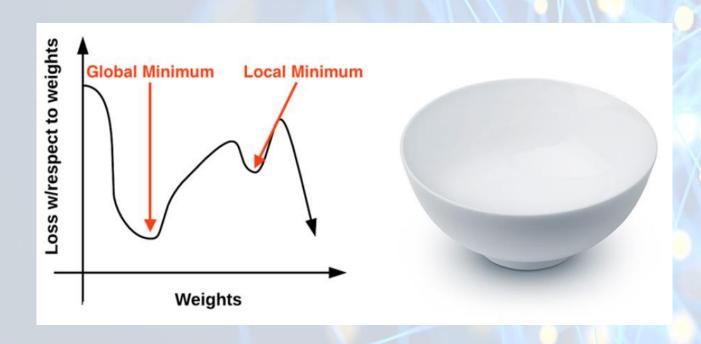


Redes Neurais Convolucionais

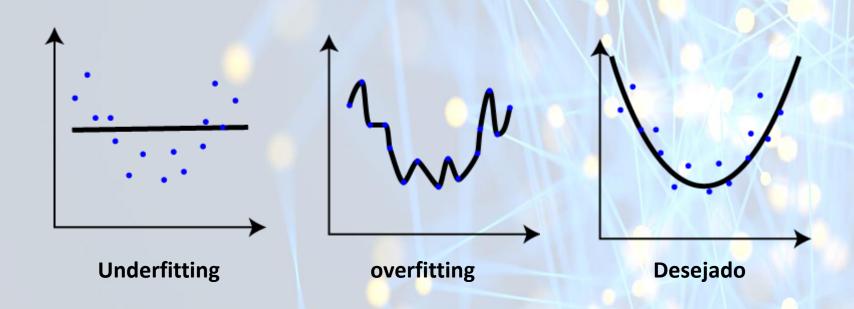


Aprendizado Supervisionado: Treinamento e Validação



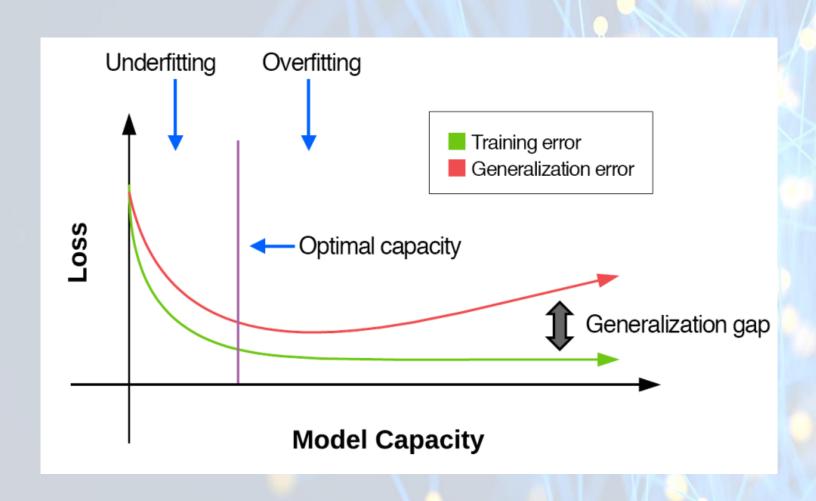


Cuidados em Machine Learning

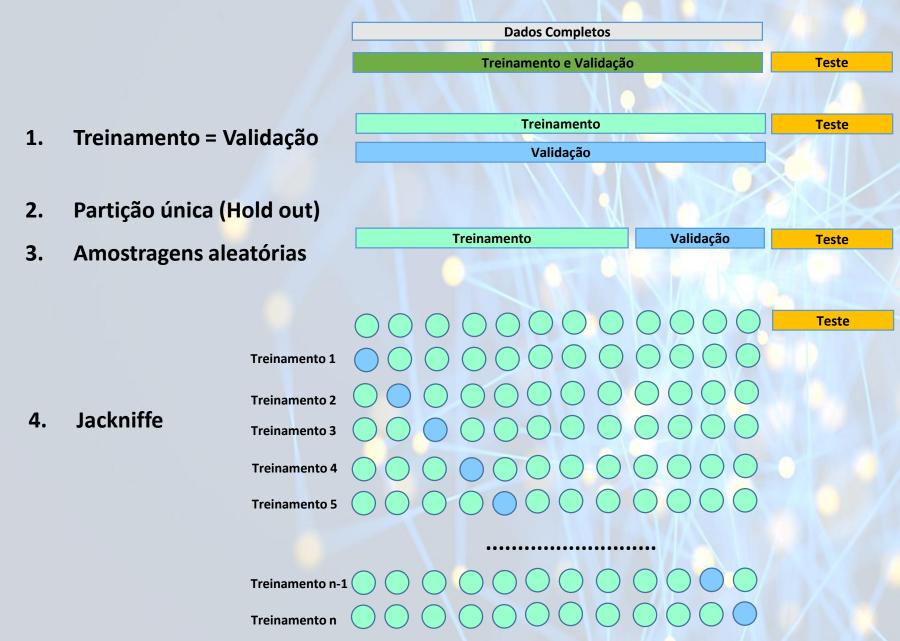


Decorar (Memorizar) não é aprender

Overfitting & Underfitting



Formas de Validação



Cross - Validação



Etapas de desenvolvimento de um sistema ML Aprendizado Supervisionado

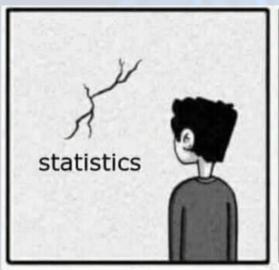
- 1. Definir o problema
- 2. Escolha das variáveis de entrada e de saída
- 3. Escolha do algoritmo de ML: Supervisionado
- 4. Definição dos hiperparâmetros do algoritmo escolhido
- 5. Treinamento, validação e teste
- 6. Uso do algoritmo testado

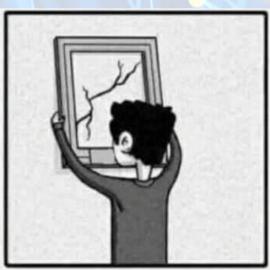
Ciência de Dados: Áreas

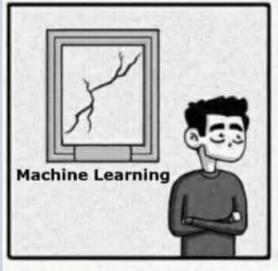


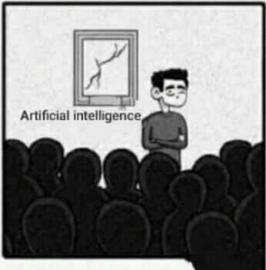


Estatística









X Estatística

Poder de Adaptação

Requer Modelagem

Não Assume Pressuposições

Pressuposições

Preocupa-se com resultado

Preocupa-se com Parâmetros





Estatística

Qual escolher?

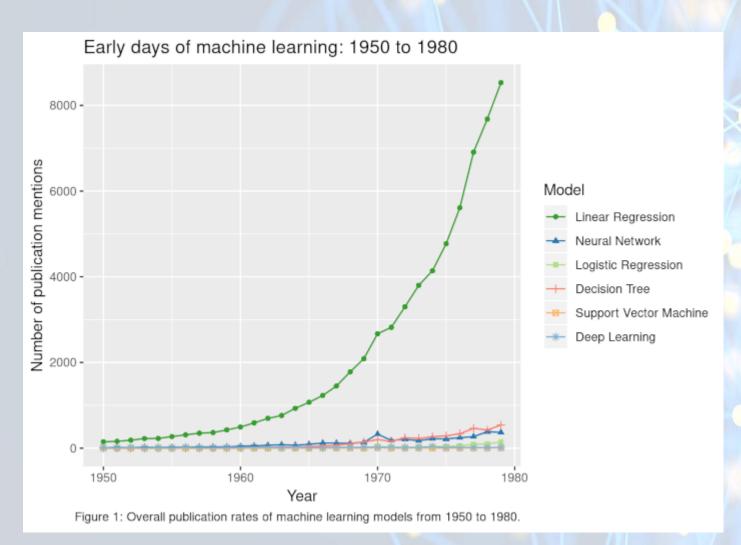
O que define essa escolha?

Qual mais eficiente?



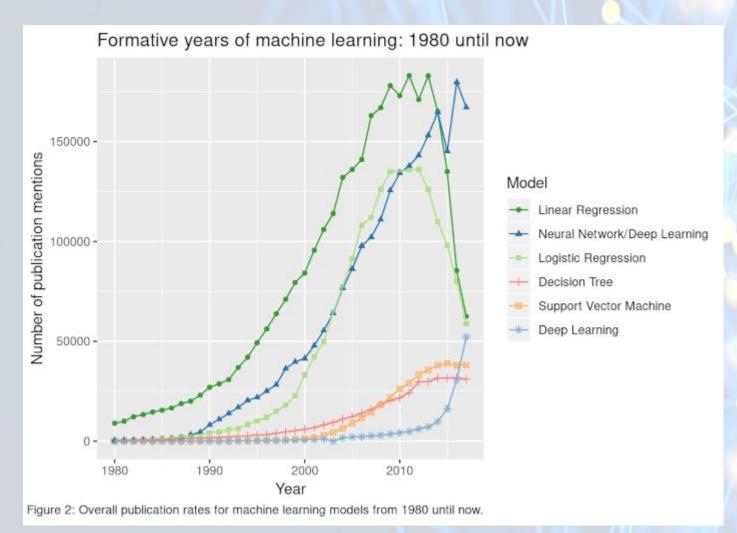


Estatística



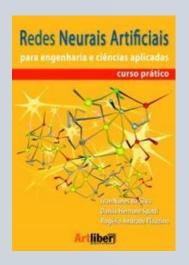


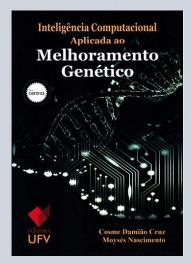
Estatística

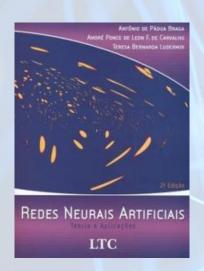


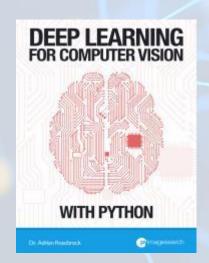
Inteligência Artificial Caminho das Pedras

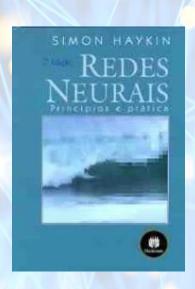
1. Leitura













Inteligência Artificial Caminho das Pedras

2. Comunidade







Google kaggle

3. Escolha do Software





4. Programação









5. Praticar







6. Censo Crítico

"Toda técnica de análise de dados é eficiente, desde que bem utilizada."

Conhecimento nunca é demais



"Não importa o quão devagar você vá, desde que você não pare" (Confúcio)