



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
PROJETO DE INICIAÇÃO CIENTÍFICA
UNIDADE ACADÊMICA DE ENGENHARIA ELÉTRICA

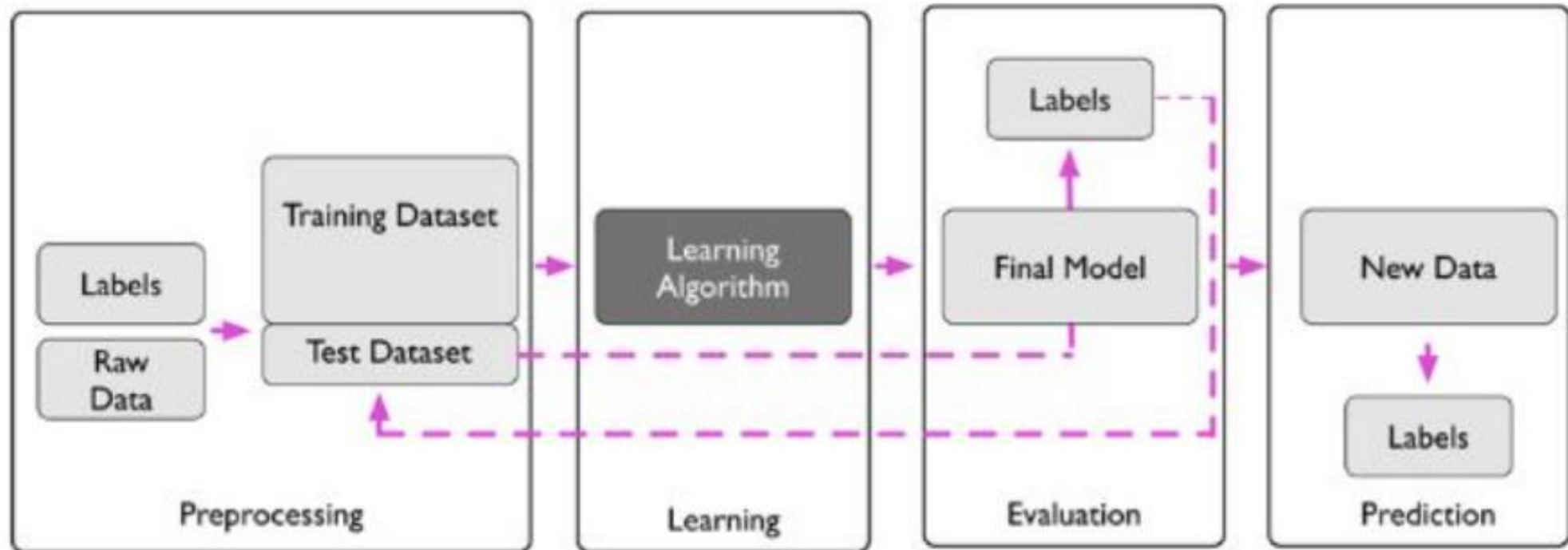
Classificação de distúrbios pulmonares em radiografias de tórax usando Redes Convolucionais

Reunião 25/09/2020

Aprendendo com os Dados

- Vivemos em uma era **Big Data**, em que o meio digital está repleto de dados úteis disponíveis para análise.
- Transformar uma grande quantidade de dados em conhecimento.
- Encontrar **padrões** e fazer **predições** sobre estados futuros.

Ciclo de vida de um algoritmo (IA)



Métodos Preditivos

- O método preditivo, como o próprio nome entrega, trata-se da ideia de fazer alguma **previsão** sobre um evento futuro, baseado em um **conjunto de atributos**.

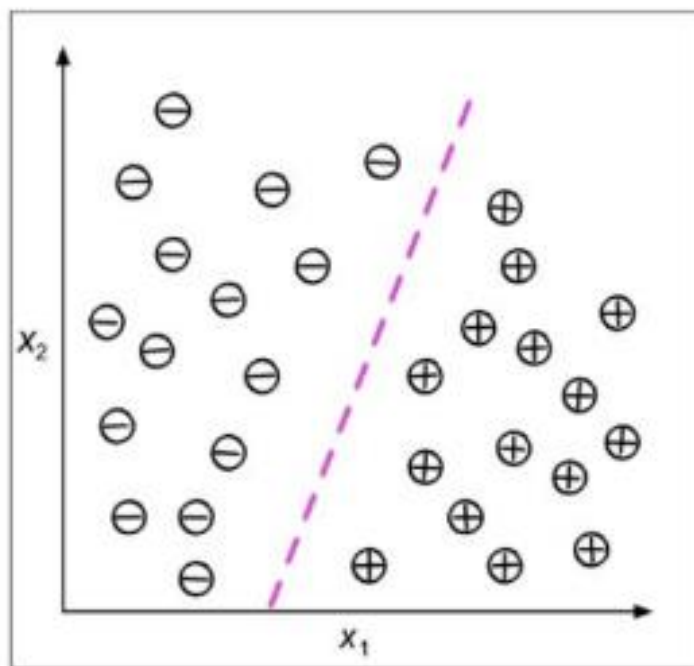
Métodos Descritivos

- O nome desse método também é bastante revelador, trata-se da ideia de **descrever** determinados objetos de acordo com os seus **comportamentos**.

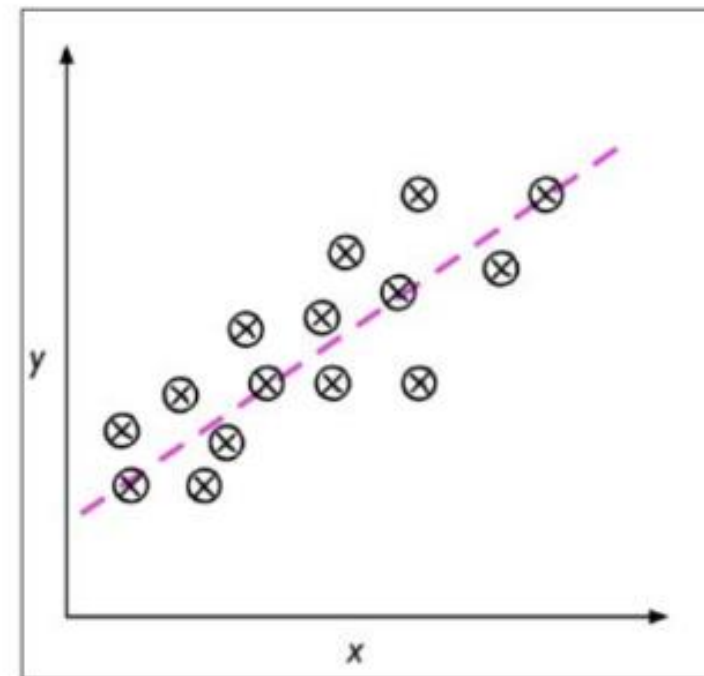
Métodos Preditivos e Descritivos

Métodos Preditivos	Métodos Descritivos
Método de Classificação	Método de Associação
Método de Regressão	Método de Agrupamento
	Método de Detecção de Desvios
	Método de Padrões Sequenciais
	Método de Sumarização

Métodos Preditivos



Classificação

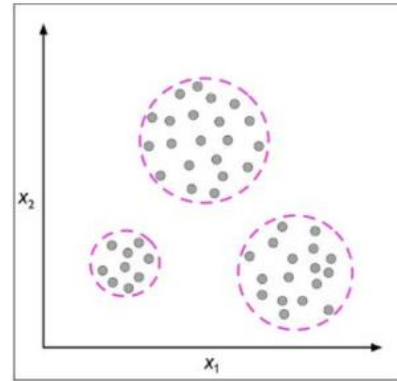


Regressão

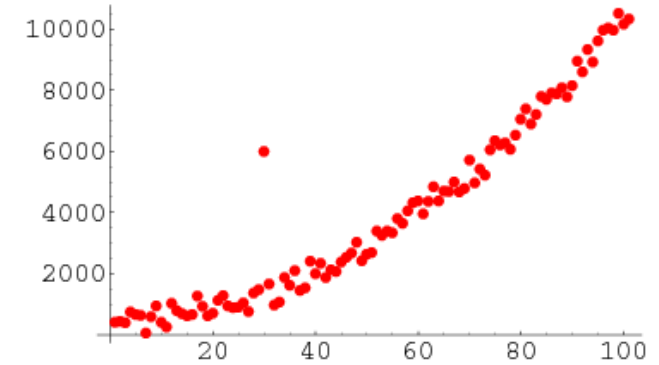
Métodos Descritivos



Associação



Agrupamento



Desvios (Outliers)



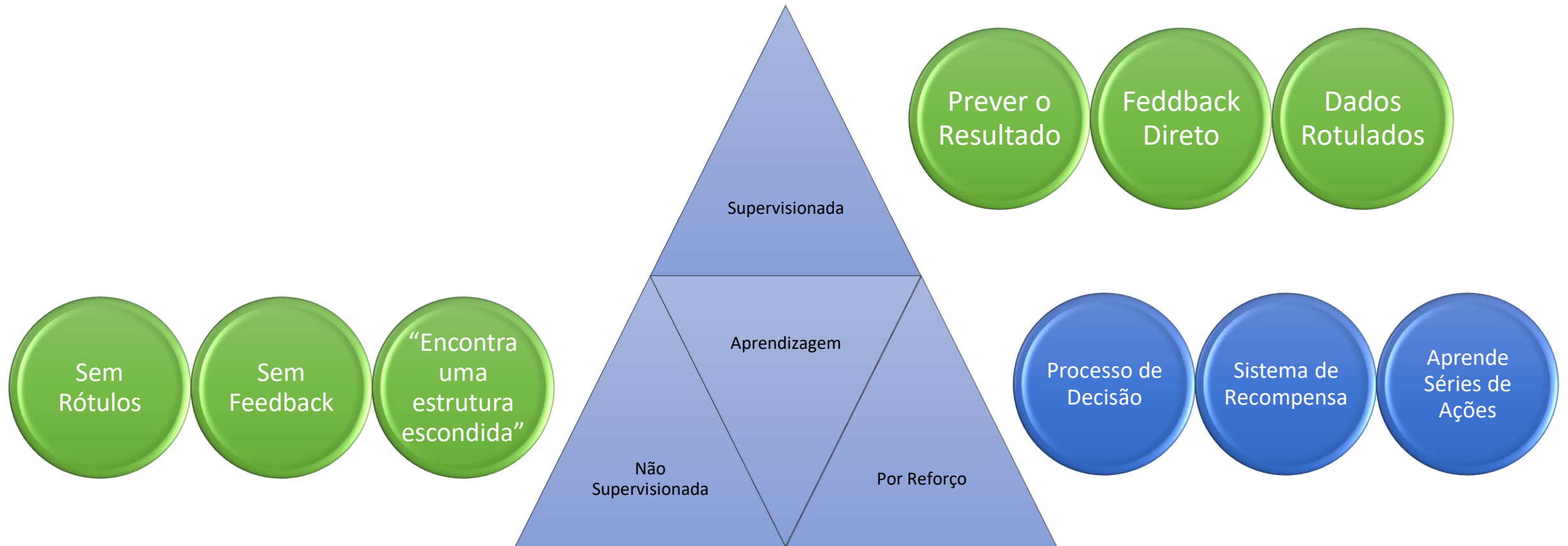
Padrões Sequenciais

Input: Article 1st sentence	Model-written headline
metro-goldwyn-mayer reported a third-quarter net loss of dlrs 16 million due mainly to the effect of accounting rules adopted this year	mgm reports 16 million net loss on higher revenue
starting from july 1, the island province of hainan in southern china will implement strict market access control on all incoming livestock and animal products to prevent the possible spread of epidemic diseases	hainan to curb spread of diseases
australian wine exports hit a record 52.1 million liters worth 260 million dollars (143 million us) in september, the government statistics office reported on monday	australian wine exports hit record high in september

Sumarização

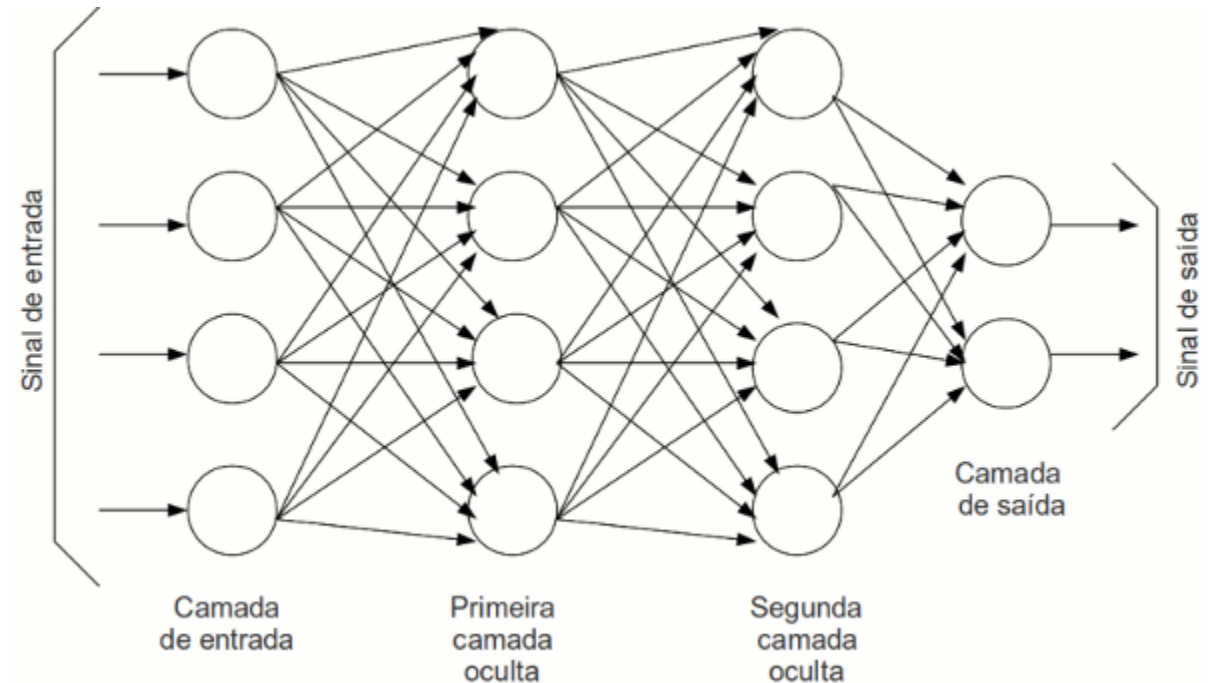
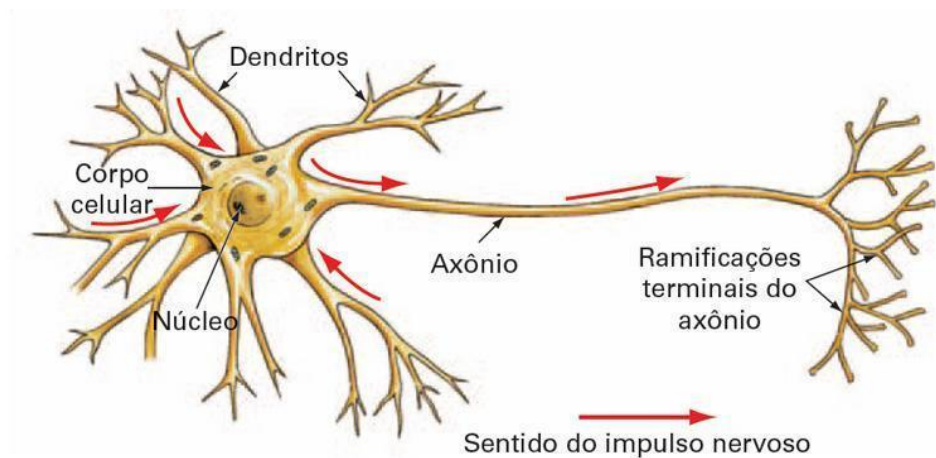
Aprendizagem Bayesiana	Algoritmo Apriori
Árvores de Decisão	Algoritmo ECLAT
Aprendizagem por Regras	Algoritmo K-means
Aprendizagem Baseada em Instâncias	Agrupamento Hierárquico
Regressão Logística	Agrupamento com DBSCAN
Máquinas de Vetores de Suporte	Redução de Dimensionalidade
Redes Neurais Artificiais	Detecção de Outliers
Regressão Linear	Outros tipos de regressões

Tipos de Aprendizagem



O que é uma Rede Neural?

- É um algoritmo de aprendizagem poderoso inspirado em como o cérebro funciona.



Dados estruturados e não estruturados

Salário	Anos de estudo	Sexo
1595.11	16	F
1774.81	12	F
886.38	16	M
2473.64	13	M
2059.46	15	F
1561.84	14	F
1850.95	17	M
2106.48	17	F
2411.22	13	M
1855.68	16	F
1999.50	17	F
2020.34	14	M
2081.46	17	M
2087.69	12	M
1522.40	14	M
2114.97	15	F
1615.36	18	F
2094.69	10	F
615.51	17	F
1874.00	15	M

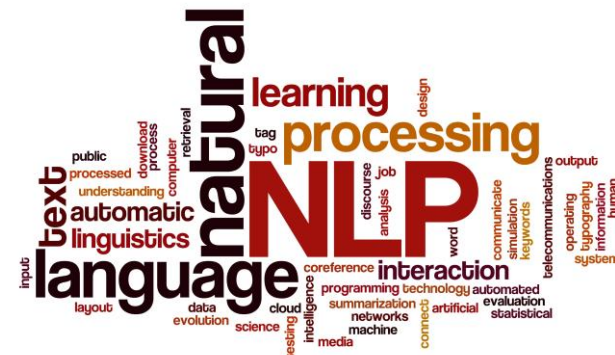
Tabela CSV



Imagens



Áudio

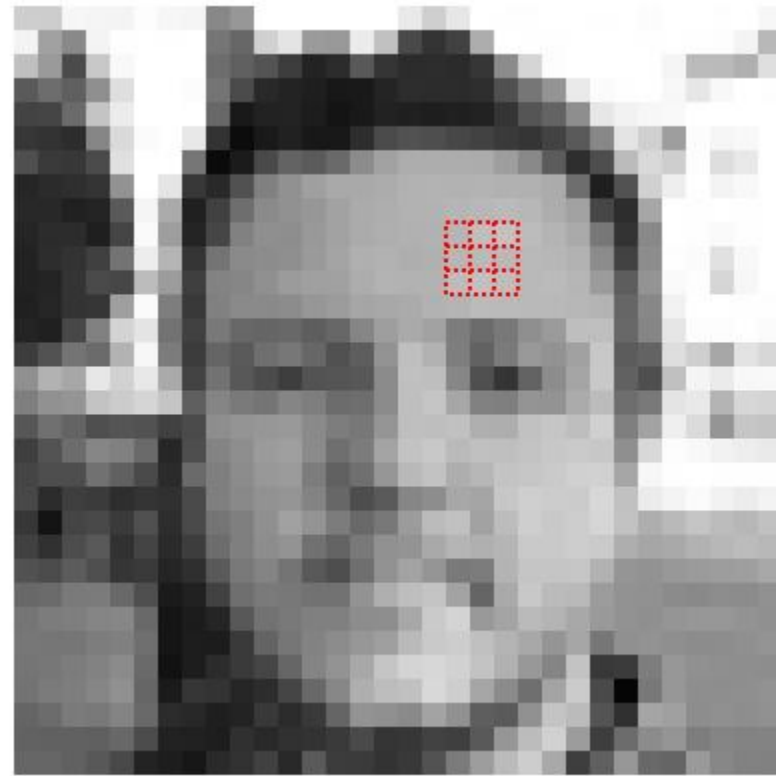


Textos

Pré-processamento dos dados

- Correção de valores faltantes;
- Correção de valores incoerentes;
- **Tratamento em imagens;**
- Aumento ou diminuição dos dados;

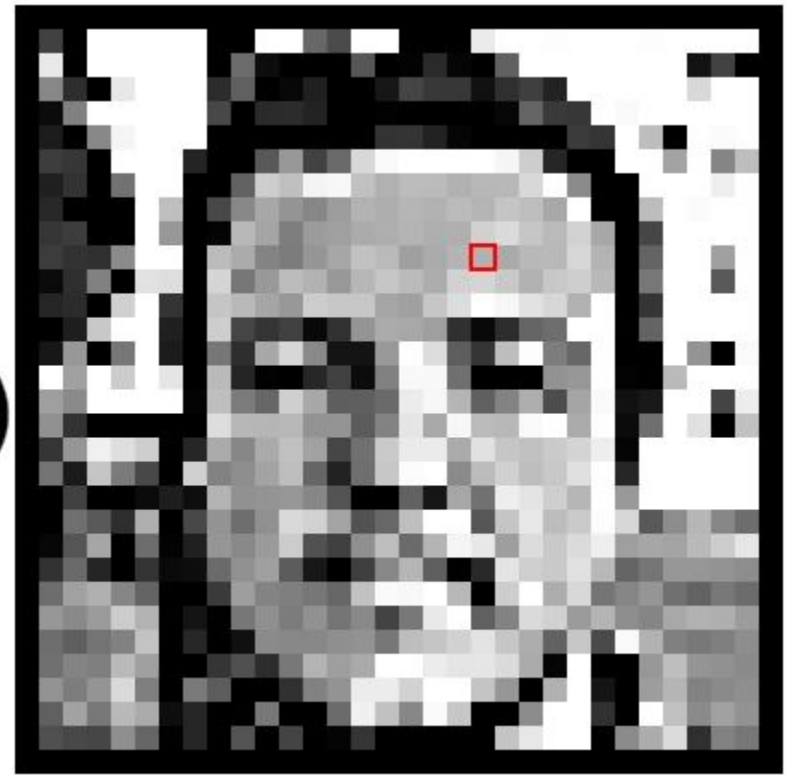
Image Kernels



input image

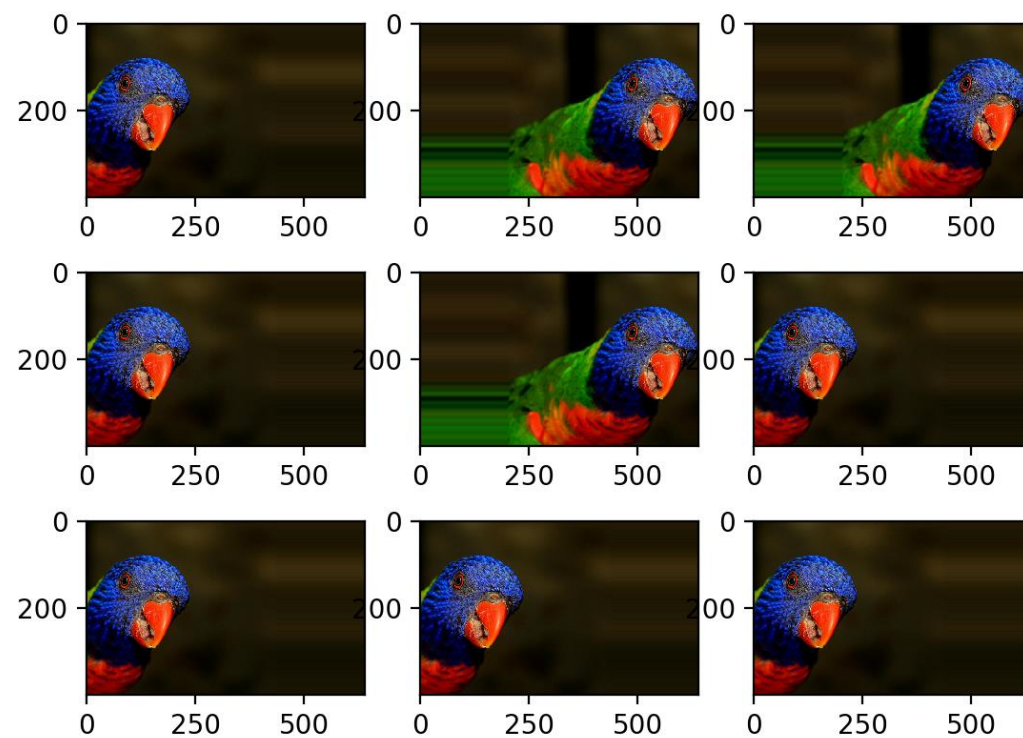
$$\left(\begin{array}{ccc} 182 & + & 191 & + & 194 \\ \times 0 & & \times -1 & & \times 0 \\ + & 183 & + & 186 & + & 185 \\ \times -1 & & \times 5 & & \times -1 \\ + & 189 & + & 193 & + & 190 \\ \times 0 & & \times -1 & & \times 0 \end{array} \right)$$
$$= 178$$

kernel:
sharpen ▼



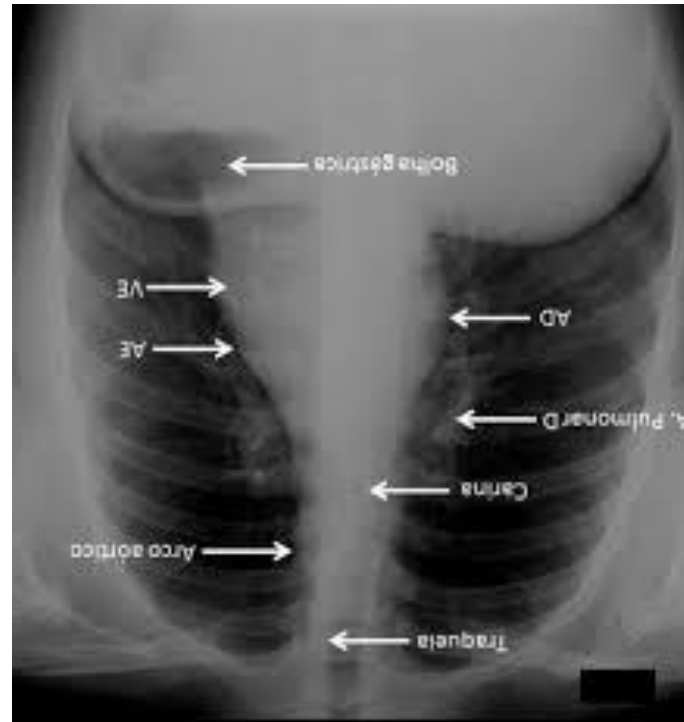
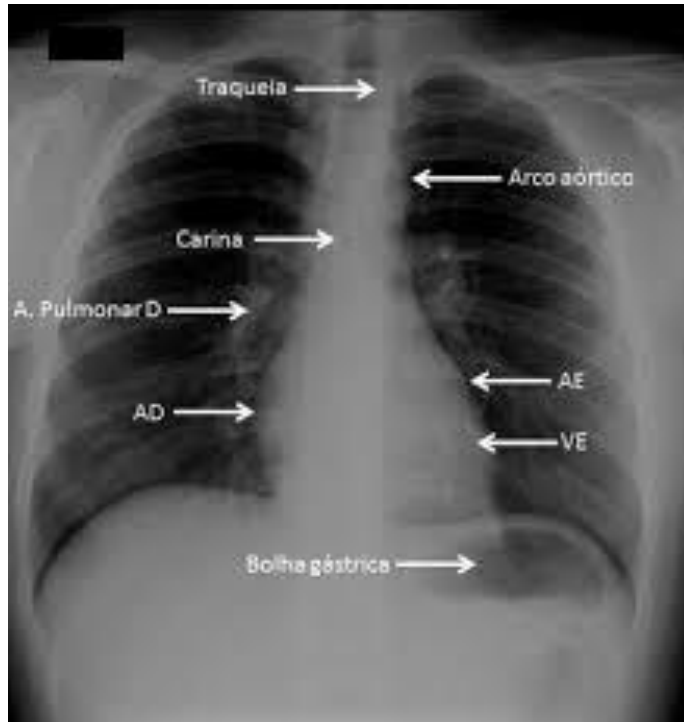
output image

Data Augmentation

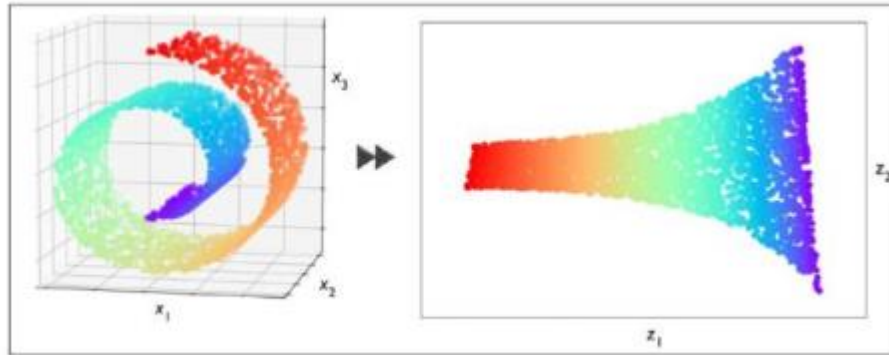


Data Augmentation

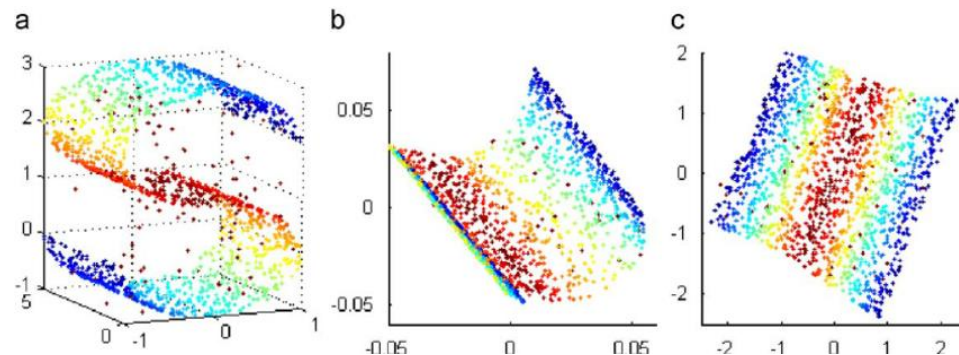
- Isso está certo?



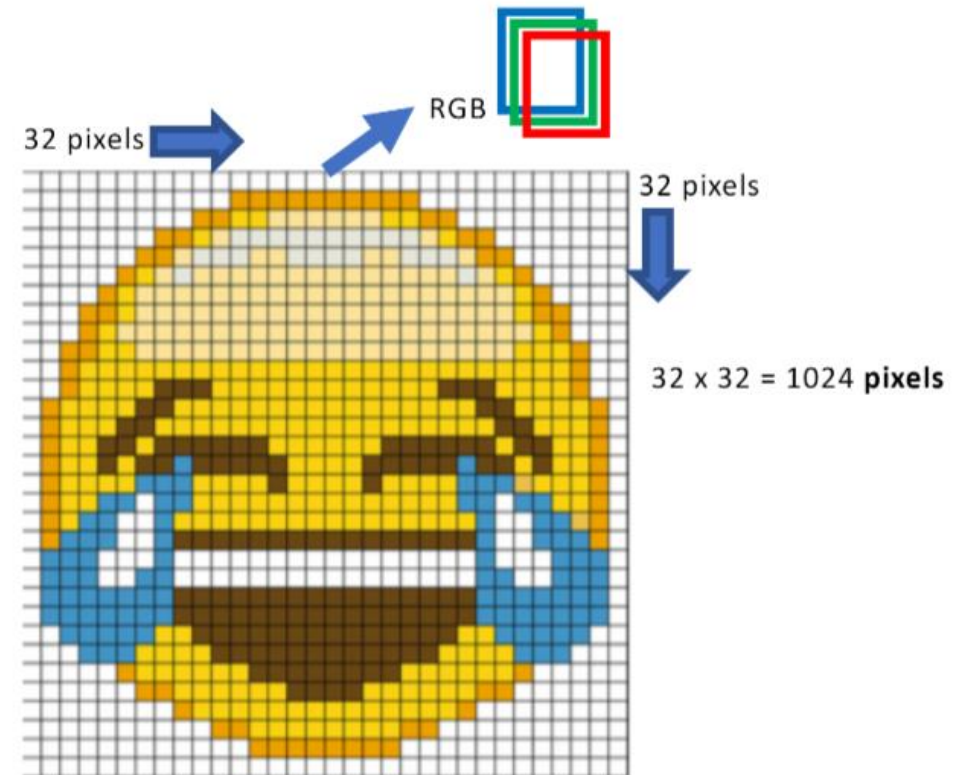
Redução da Dimensionalidade



3D – 2D

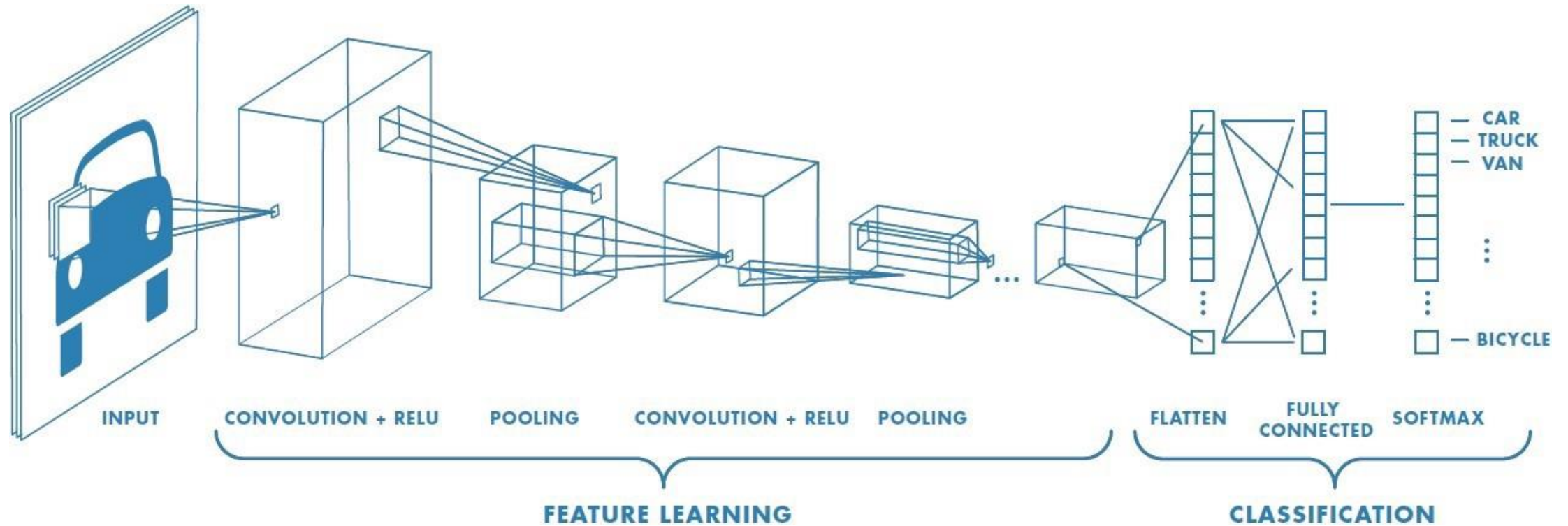


3D – 2D



Imagens Digital

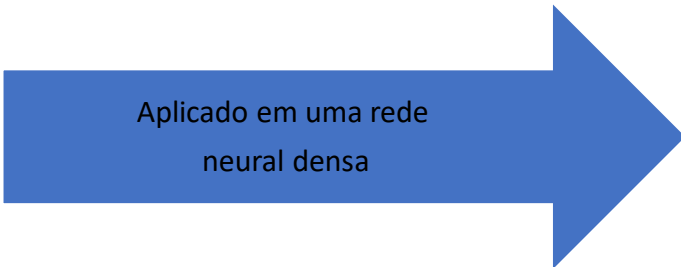
Etapas de uma Rede Neural Convolucional



Etapas de uma Rede Neural Convolutacional

- Usado para visão computacional.
- Detecção de características importantes na imagem.

Aplicado em uma rede
neural densa



Etapas

1 – **Operador de Convolução**: adicionar cada elemento da imagem para seus vizinhos, ponderando por um kernel (Construção de um mapa de características através de um detector de características).

- > **Perca de informação e melhoria no processamento.**

2 – **Pooling**: Seleciona as características mais relevantes.

- > **Reduz overfitting e ruídos desnecessários.**

3 - **Flattening**: converter os dados em uma matriz unidimensional para inseri-los na próxima camada. Achatamos a saída das camadas convolucionais para criar um único vetor.

- > **Melhoria no Processamento.**

Vetorização em Python

```
[9]: import numpy as np
import time

# definindo vetores aleatórias
a = np.random.rand(10000000)
b = np.random.rand(10000000)

tic = time.time()

# multiplicando os vetores
c = np.dot(a, b)

toc = time.time()

print("Vectorized Version: " + str(1000*(toc-tic)) + " ms")
```

Vectorized Version: 10.128259658813477 ms

```
[10]: import numpy as np
import time

# definindo vetores aleatórias
a = np.random.rand(10000000)
b = np.random.rand(10000000)

c = 0

tic = time.time()

# usando um loop para realizar a multiplicação dos vetores
for i in range(10000000):
    c += a[i]*b[i]

toc = time.time()

print("Loop Version: " + str(1000*(toc-tic)) + " ms")
```

Loop Version: 14074.564218521118 ms

Broadcasting Python

O termo broadcasting descreve como o numpy trata matrizes com diferentes formas durante as operações aritméticas.

```
[6]: B = np.array([1, 2, 3, 4]) + 100
```

```
print(B)
```

```
[101 102 103 104]
```

```
[7]: C = np.array([[1, 2, 3, 4],  
                  [5, 6, 7, 8]]) * B
```

```
print(C)
```

```
[[101 204 309 416]  
 [505 612 721 832]]
```

Referências

- *Lago, S. P. **Introdução a Inteligência Artificial**, USP, 2020.*
- *Granatyr, J. **IA Expert**. Disponível em: <https://iaexpert.com.br/>. Acesso em 27 abr. 2020.*
- RASCHKA, Sebastian; MIRJALILI, Vahid. **Python Machine Learning**, 2nd ED. Packt Publishing, 2017.
- NG, Andrew. **Neural networks and deep learning**. Disponível em: <https://www.coursera.org/learn/neural-networks-deep-learning?specialization=deep-learning>. Acesso em 24 set. 2020.