



Machine Learning com R

Seja bem-vindo!!!

CLASSIFICAÇÃO

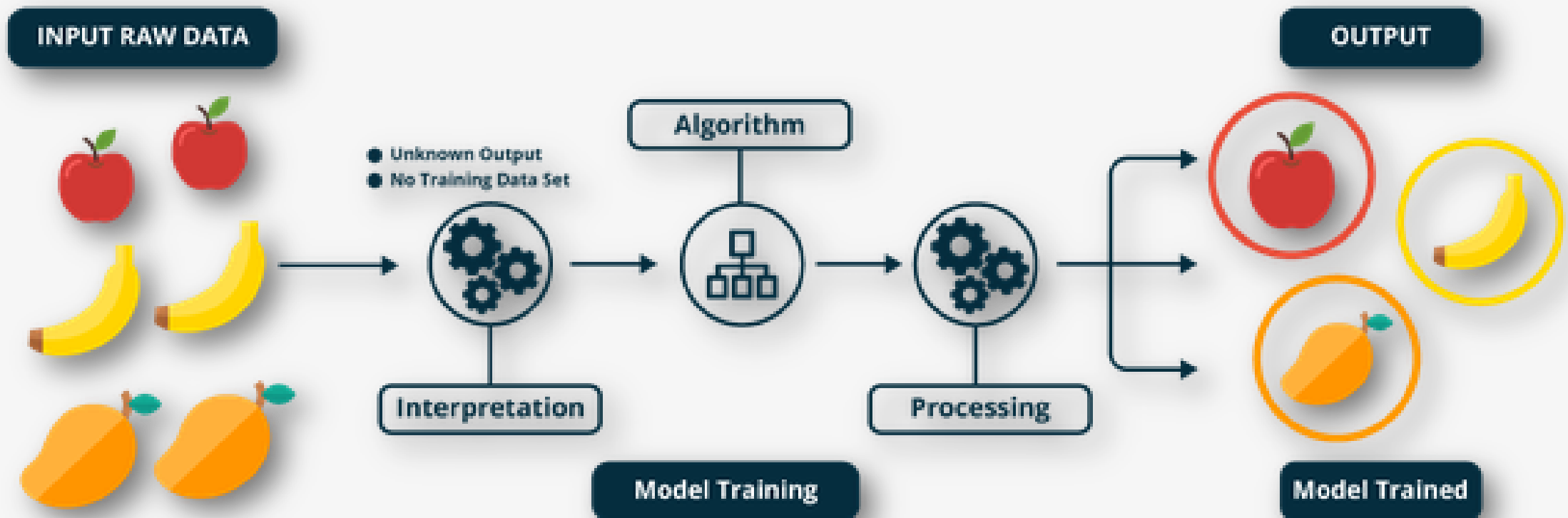
Machine Learning

- **Classificação:**

- É um técnica chamada de APRENDIZADO SUPERVISIONADO;
- O modelo aprende a partir de uma base de dados com classes (target);
- Pode ser:
 - **binária** : 0/1, S/N, etc.
 - **multi-classe**: 0/1/2, manhã/tarde/noite, sol/chuva/ventania etc.

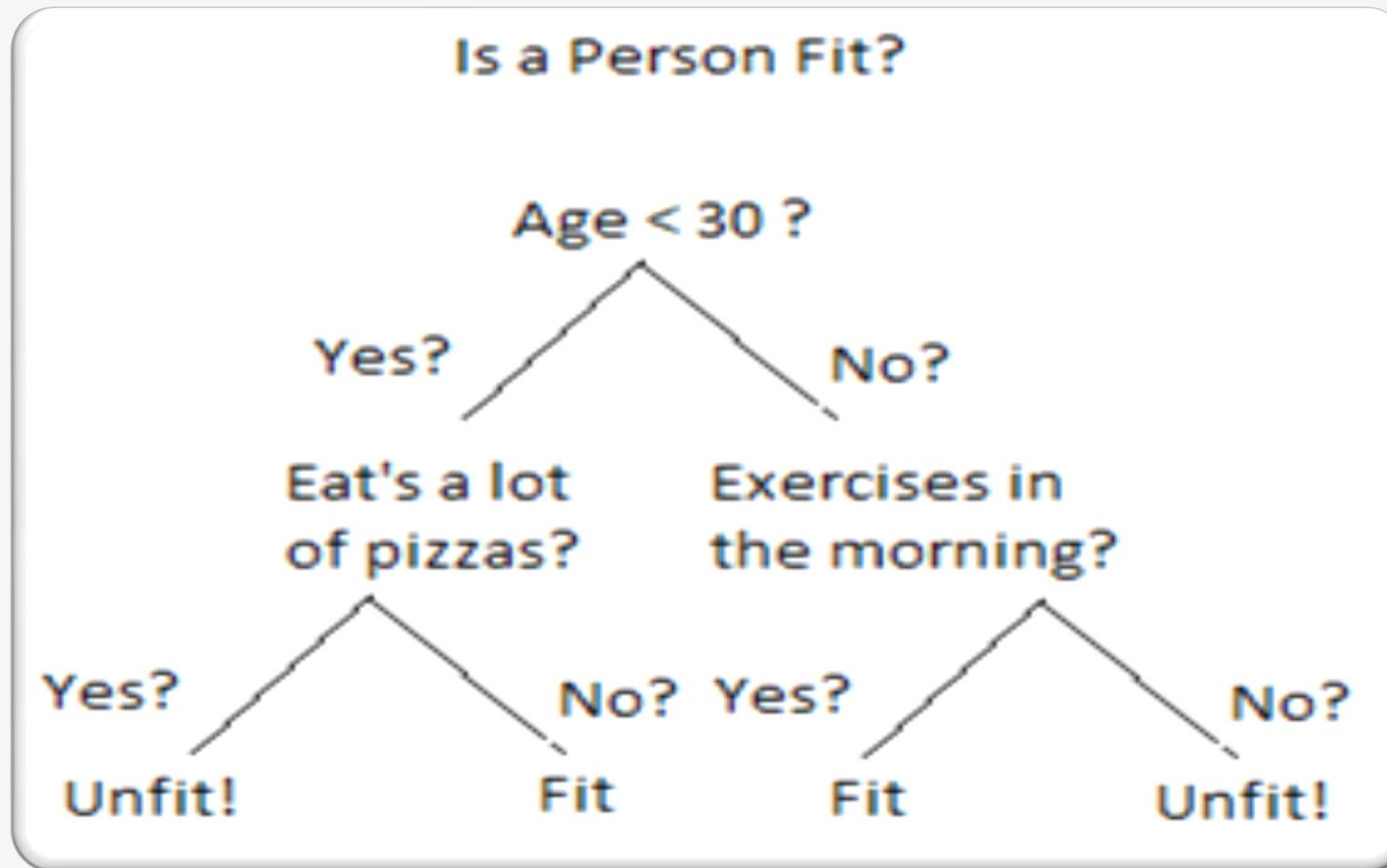
Machine Learning

- Classificação:



Machine Learning

- Classificação:



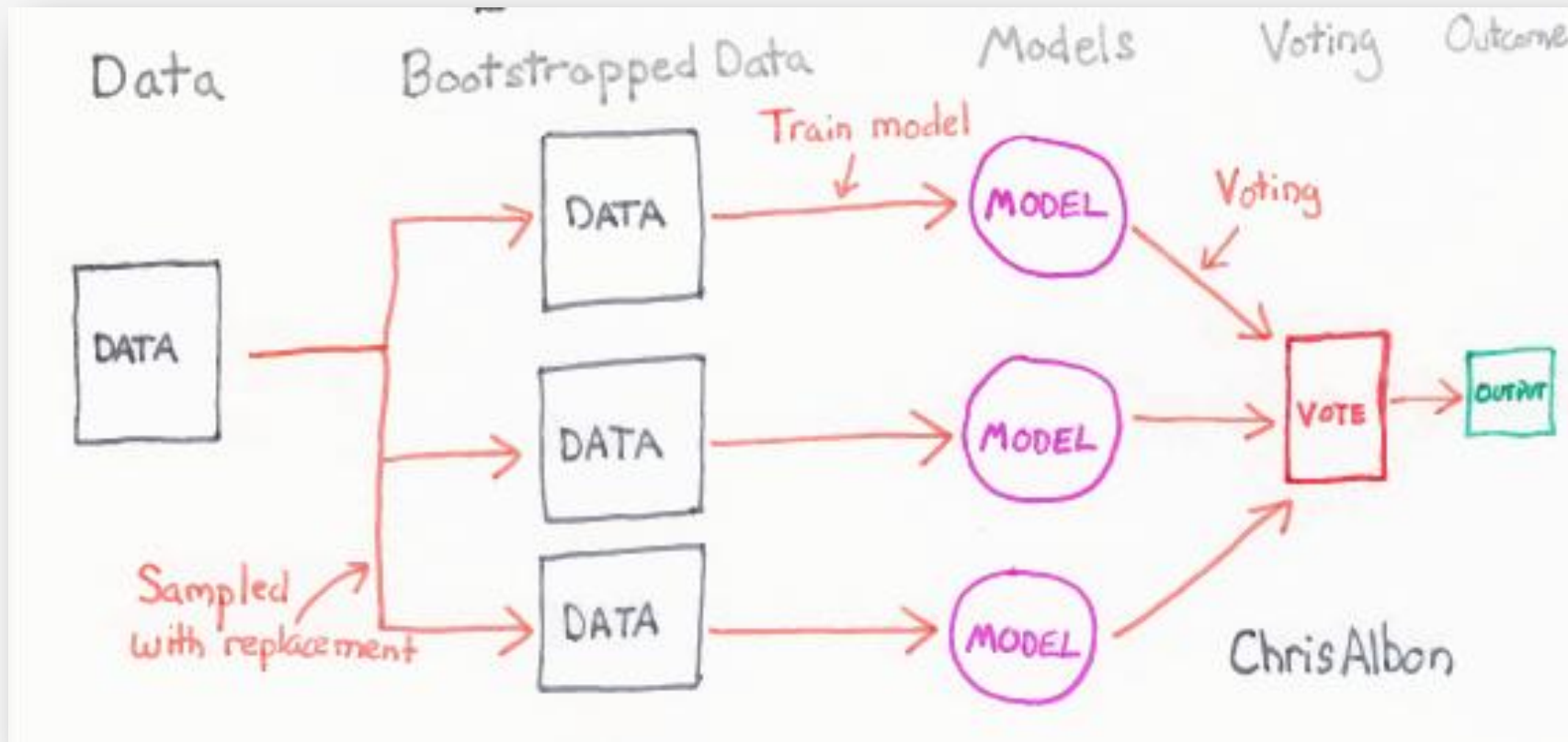
Machine Learning

- **Algoritmos de classificação:**

- SVM
- DECISION TREES
- LOGISTIC REGRESSION
- NAIVE BAYES
- **RANDOM FOREST**
- ETC

Machine Learning

- Random Forest



Machine Learning

- **Random Forest :**

- É um algoritmo que executa árvore de decisão com mais de um tipo de algoritmo de classificação;
- Faz uma combinação dos modelos para obter um melhor resultado final;
- Ao final faz uma “votação” pra escolher o melhor resultado;

Machine Learning

- **Avaliando o modelo de classificação**
 - Verificar o quanto o modelo está acertando/prevendo os resultados;
 - Pode ser utilizado: **Matrix Confusion, Accuracy, recall, precision, fscore.**

Machine Learning

- **Accuracy**

- Métrica para avaliar os modelos, da proporção dos resultados rotulados

corretamente;

- 0% - 100%

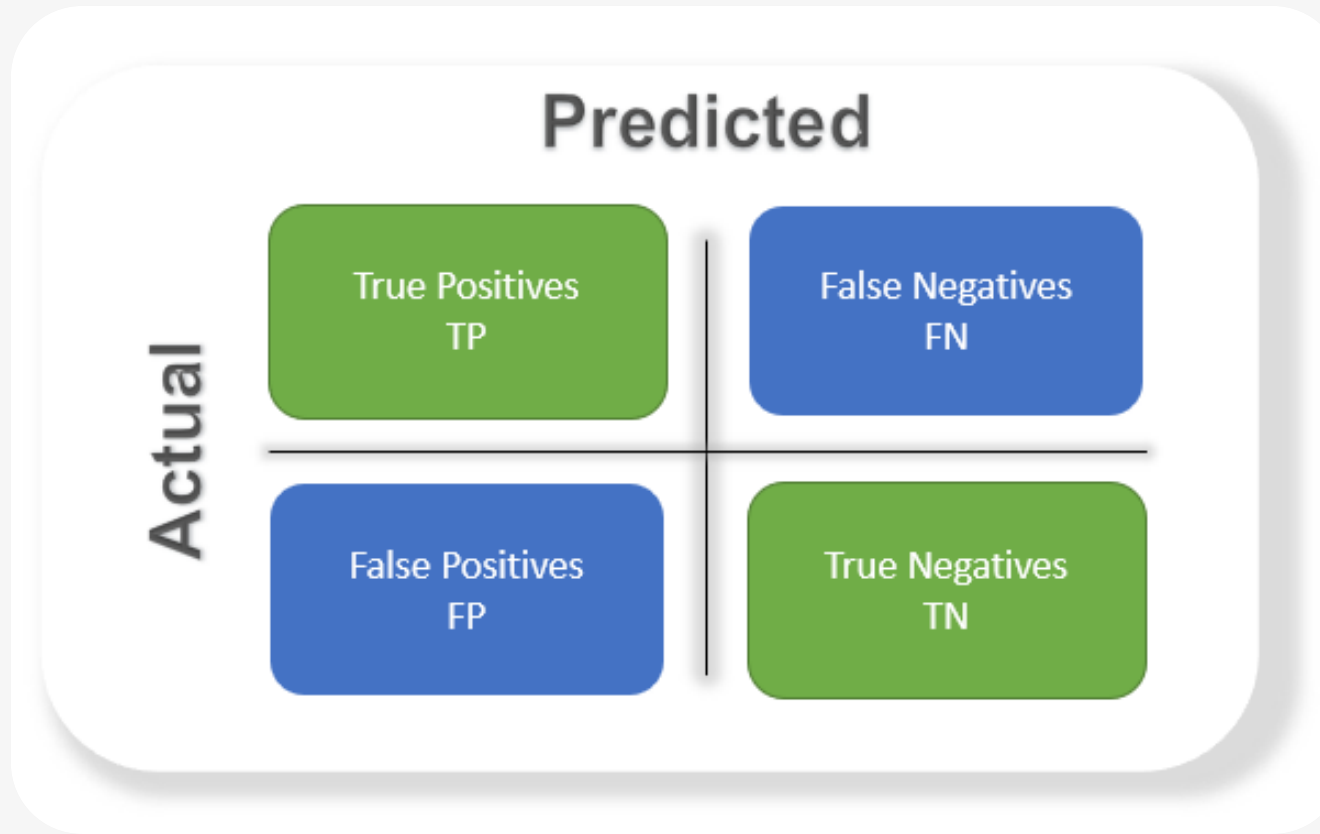
$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

- **PODE SER UM PROBLEMA, PORQUE NÃO REPRESENTA A REAL PERFORMANCE**

DO MODELO DE ML

Machine Learning

- **Matrix Confusion (matriz de confusão):** Representação dos resultados obtidos a partir do modelos de ML e os valores reais.



Machine Learning

- **Matrix Confusion (matriz de confusão)**

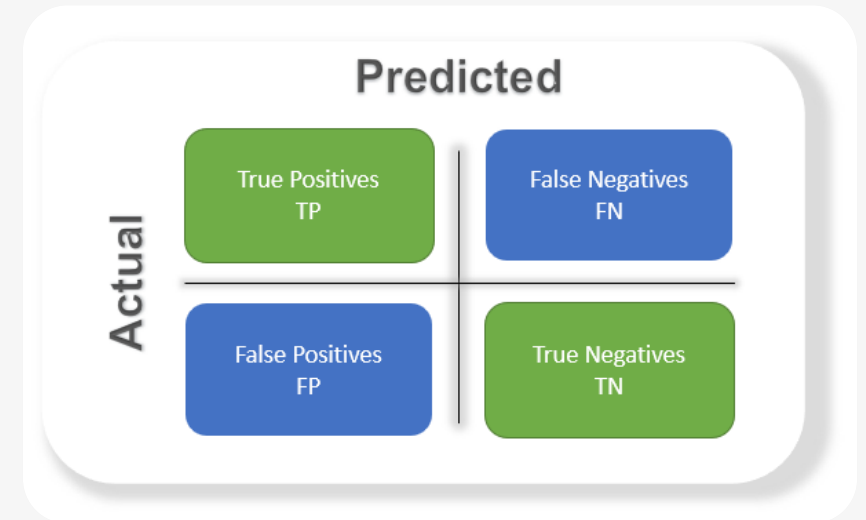
- Representação dos resultados obtidos a partir do modelos de ML e os valores reais.

- **True Positives:** valores classificados corretamente para classe desejada

- **True Negatives:** valores classificados corretamente para a classe oposta da desejada.

- **False Negatives:** valores classificados incorretamente para a classe oposta da desejada.

- **False Positives:** valores classificados para a classe desejada incorretamente.



Machine Learning

- Matrix Confusion (matriz de confusão)

- In [2]:

```
# 1 para grávida, 0 para não grávida  
valores_reais      = [1, 0, 1, 0, 0, 0, 1, 0, 1, 0]  
valores_preditos   = [1, 0, 0, 1, 0, 0, 1, 1, 1, 0]
```

		Valores preditos	
		Grávida	Não Grávida
Valores reais	Grávida	3	1
	Não grávida	2	4

Machine Learning

- **Recall:**

- Qual proporção de positivos foi identificados corretamente;
- Quão bom o modelo é para prever positivos.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Machine Learning

- **Recall:**

- Qual proporção de positivos foi identificados corretamente;
- Quão bom o modelo é para prever positivos.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

Machine Learning

- **Precision:**

- Proporção de predições positivas realmente corretas.
- Quanto o modelo trabalhou bem

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Machine Learning

- **Precision:**

- Proporção de predições positivas realmente corretas.
- Quanto o modelo trabalhou bem

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

Machine Learning

- **FSCORE (F1):**

- Mostra o balanço entre a precision e o recall do modelo.

$$2 * \frac{precision * recall}{precision + recall}$$

Machine Learning

- **Random Forest :**

- É um algoritmo que executa árvore de decisão com mais de um tipo de algoritmo de classificação;
- Faz uma combinação dos modelos para obter um melhor resultado final;
- Ao final faz uma “votação” pra escolher o melhor resultado;

Machine Learning



Machine Learning

•Referências:

- <https://towardsdatascience.com/supervised-machine-learning-classification-5e685fe18a6d>
- <https://gabrielshade.github.io/2019/03/12/ml-classificacao-metricas.html>
- <https://blog.idwall.co/o-que-e-machine-learning/>
- <https://paulovasconcellos.com.br/como-saber-se-seu-modelo-de-machine-learning-est%C3%A1-funcionando-mesmo-a5892f6468b>
- <https://towardsdatascience.com/supervised-machine-learning-classification-5e685fe18a6d>
- <https://medium.com/datadriveninvestor/classification-in-machine-learning-db33514c77ad>