



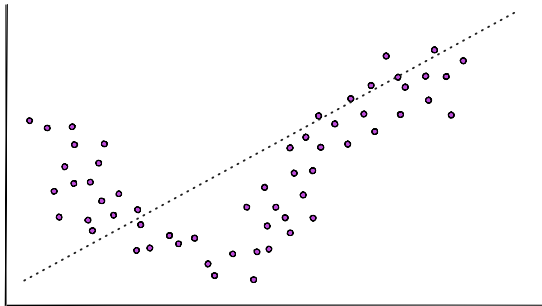
Pq avaliar?

Unitins - Sistemas de Informação
Inteligência Artificial
Professor Marco Antonio Firmino de Sousa

Pq avaliar?

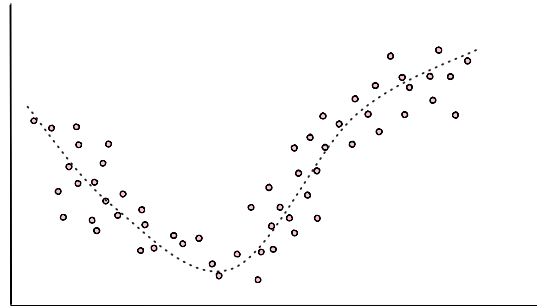
Generalização e sobre ajuste

Underfitted



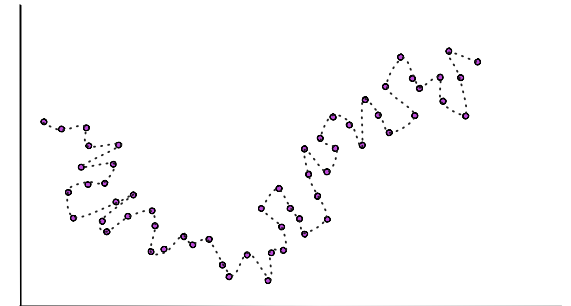
subajuste

Robust



balanceado

Overfitted



sobreajuste

Novos Dados

Organização dos dados

Conjunto de Treinamento

Utilizado para ajustar
os parâmetros do modelo

Conjunto de Validação

Usado para ajustar
hiperparâmetros
e evitar overfitting

Conjunto de Teste

Usado para avaliar o
desempenho final do modelo

→ Próxima aula

→ Validação Cruzada

Métricas de Avaliação para Classificação



```
graph TD; A[Métricas de Avaliação para Classificação] --- B[Acurácia]; A --- C[Matriz de Confusão]; A --- D[Precisão e Revocação]; A --- E[F1-Score]; A --- F[Curva ROC e AUC]
```

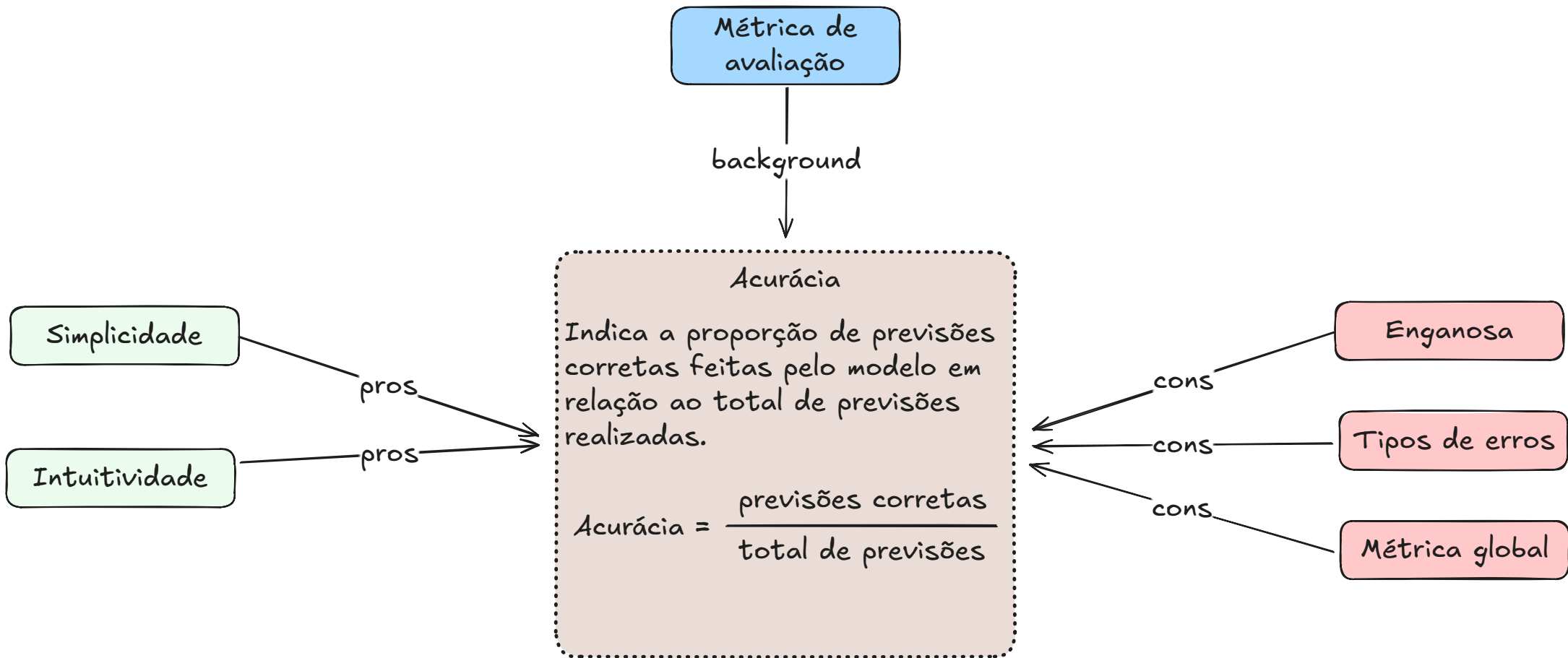
Acurácia

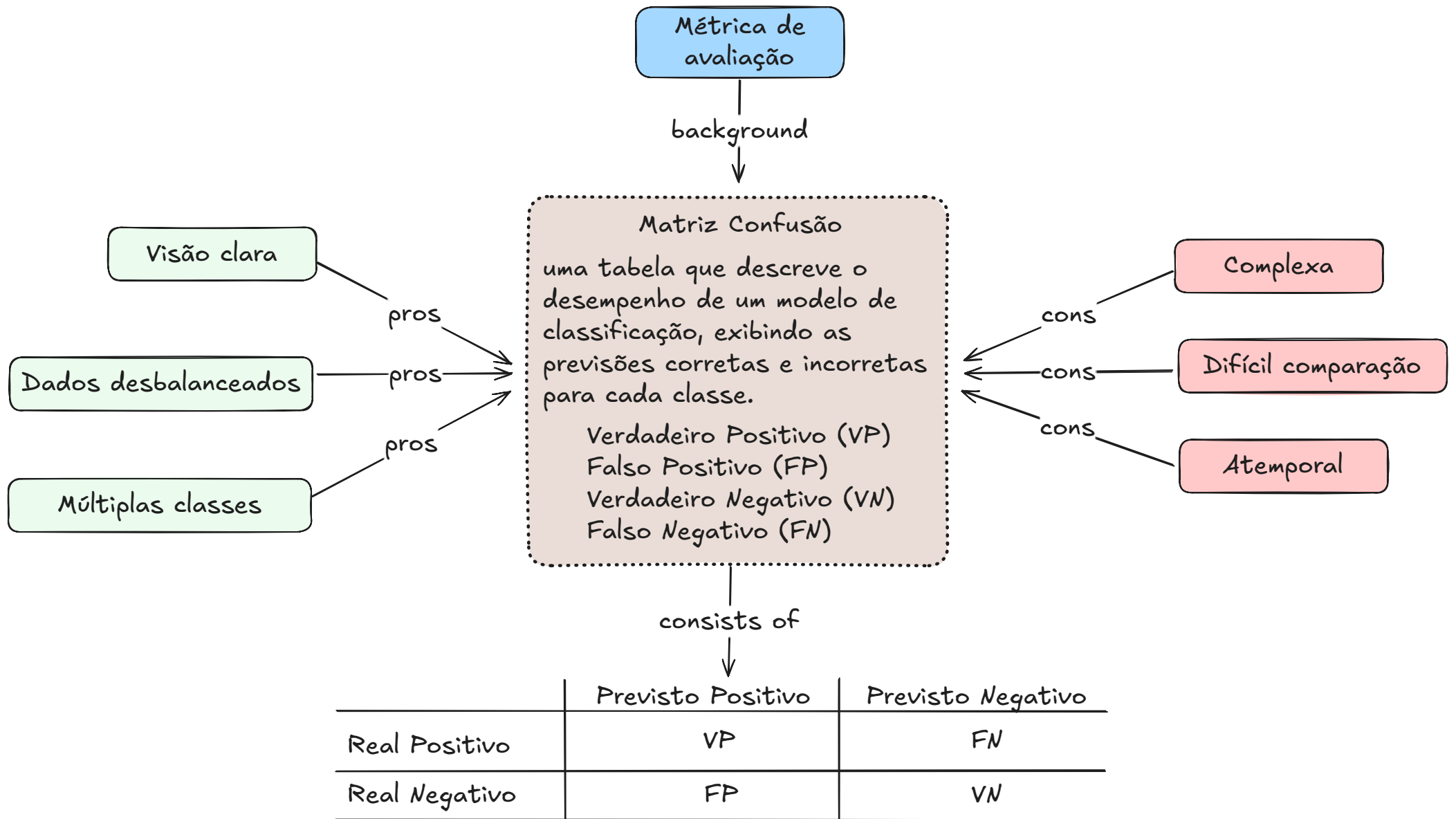
Matriz de Confusão

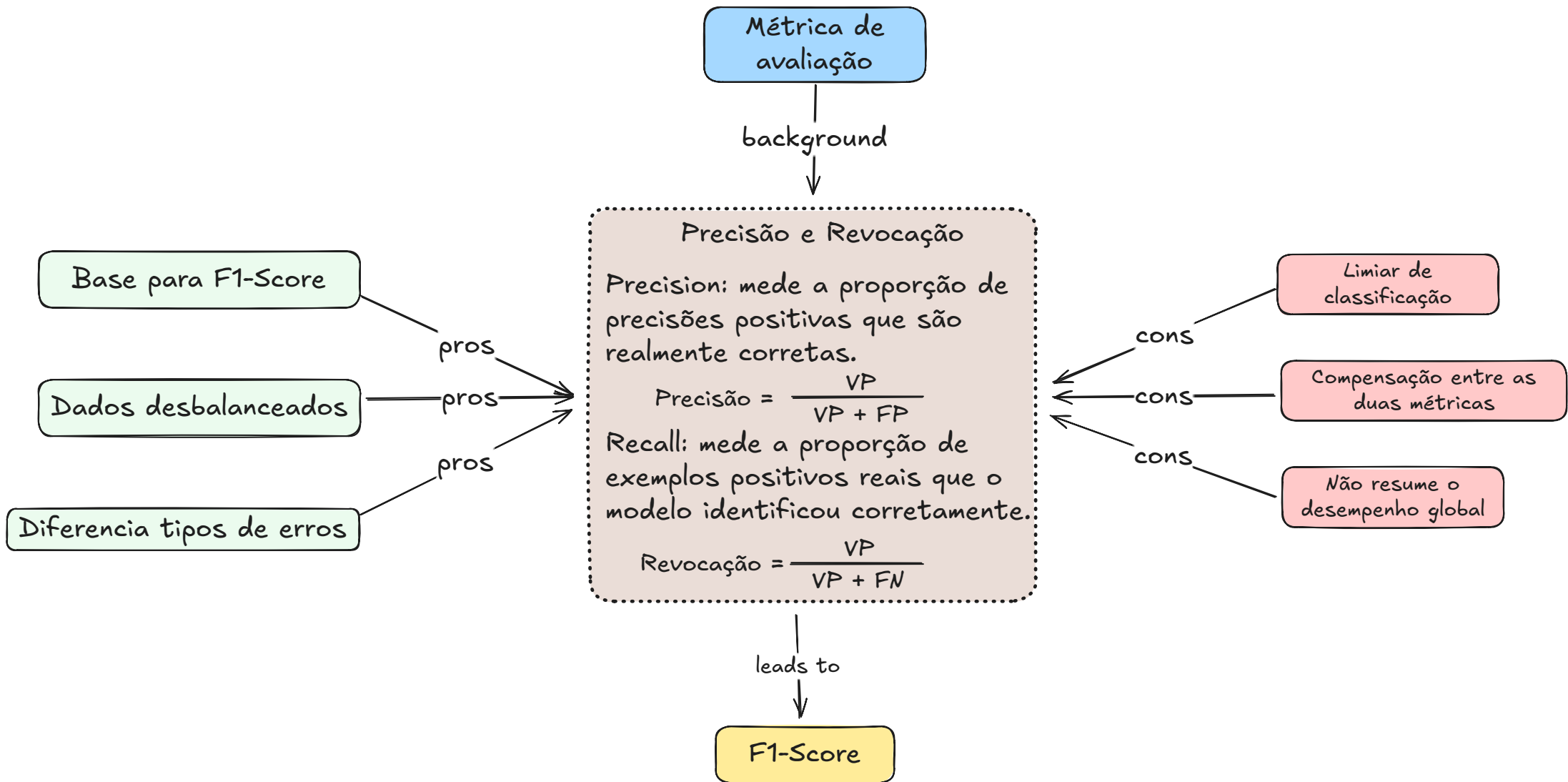
Precisão e Revocação

F1-Score

Curva ROC e AUC







Métrica de avaliação

background

F1-Score

É a média harmônica entre a Precisão e a Revocação.

$$\text{F1-Score} = 2 \times \frac{\text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}}$$

Equilíbrio entre Precision e Recall

pros

Dados desbalanceados

pros

Métrica única e intuitiva

pros

Equilibrado, mas nem sempre adequado

cons

Ignora os verdadeiros negativos

cons

Não considera a distribuição de classes

cons

Métrica de avaliação

background

Curva ROC e AUC

Curva ROC é um gráfico que exibe a relação entre a Taxa de Verdadeiros Positivos e a Taxa de Falsos Positivos.

AUC é uma métrica que resume a performance do modelo em uma única pontuação, representando a área sob a curva ROC.

Robusto para modelos binários

pros

Dados desbalanceados

pros

Comparação de modelos

pros

cons

Foco em classificação binária

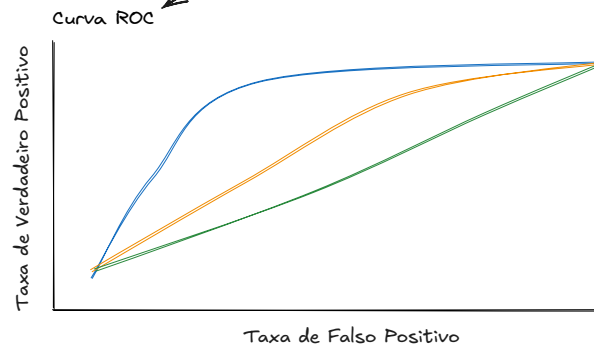
cons

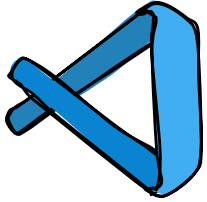
Dificuldade de interpretação

cons

Não reflete o desempenho em cenários específicos

illustrates





 [https://github.com/sousamaf/AI-Lab/
algorithms/neural_networks/mlp/mlp_torch_avaliacao.py](https://github.com/sousamaf/AI-Lab/algorithms/neural_networks/mlp/mlp_torch_avaliacao.py)