

## Métricas e Validação de Modelos de Aprendizado de Máquina

A tarefa 23 me levou a uma análise das métricas para avaliação de desempenho em modelos de aprendizado de máquina, concentrando-se em problemas de classificação. A avaliação é conduzida com base em uma série de métricas fundamentais que oferecem insights valiosos sobre a eficácia dos modelos.

### Matriz de Confusão

A matriz de confusão fornece uma visão detalhada das previsões corretas e incorretas realizadas pelo modelo para cada classe do conjunto de dados. Os elementos fundamentais da matriz incluem VN (Verdadeiro Negativo), VP (Verdadeiro Positivo), FN (Falso Negativo) e FP (Falso Positivo).

### Métricas Básicas

- **Acurácia:** Calculada como a quantidade de acertos dividida pelo total de tentativas ( $VN + VP + FN + FP$ ).
- **Precisão:** Quantidade de positivos verdadeiros dividida pela soma de positivos verdadeiros e falsos positivos ( $VP / (VP + FP)$ ).
- **Recall:** Quantidade de positivos verdadeiros dividida pela soma de positivos verdadeiros e falsos negativos ( $VP / (VP + FN)$ ).
- **Especificidade:** Calculada como a quantidade de negativos verdadeiros dividida pela soma de negativos verdadeiros e falsos positivos ( $VN / (VN + FP)$ ).

### Curva ROC

A curva ROC representa graficamente a relação entre Especificidade (X) e Sensibilidade (Y). Em suma, quanto mais próxima do canto superior esquerdo, melhor o desempenho do modelo.

### SHAP (SHapley Additive exPlanations)

O SHAP é uma ferramenta poderosa para compreender e desenvolver modelos de machine learning, oferecendo insights sobre as previsões do modelo. Recursos como `force_plot`, `summary_plot` e `dependence_plot` permitem uma análise detalhada das características e pesos do modelo.