

Prática 2 – Classificando dados do Kaggle

Objetivo

Aplicar os algoritmos de classificação [Árvore de Decisão](#), [Naive Bayes](#), [SVM](#), [kNN](#) na resolução de um problema real de predição, por meio da participação na competição "Santander Customer Transaction Prediction" no Kaggle. A atividade tem como foco o desenvolvimento prático, a comparação entre modelos e a interpretação crítica dos resultados obtidos.

Descrição da Atividade

Utilize os algoritmos mencionados para desenvolver soluções para a competição:

- <https://www.kaggle.com/competitions/santander-customer-transaction-prediction/overview>

Como apoio, consulte o notebook apresentado em sala de aula (disponível na pasta *da prática* no Google Drive).

Requisitos da Atividade

- Construa modelos utilizando os três algoritmos: [Árvore de Decisão](#), [Naive Bayes](#) e [kNN](#);
- Realize avaliação local utilizando divisão entre treino e teste (*train_test_split*);
- Submeta as predições ao Kaggle e registre os **scores obtidos no Leaderboard**;
- **Tente melhorar o desempenho dos modelos** por meio do **ajuste dos hiperparâmetros** (por exemplo: profundidade da árvore, número de vizinhos no KNN etc.);
- Compare os resultados locais (avaliação com treino/teste) com os obtidos no Kaggle e analise possíveis diferenças (como *overfitting* ou *underfitting*). Para isso, no fim do notebook,:
 - Crie uma tabela como:

Algoritmo	Avaliação Local (antes)	Kaggle (antes)	Avaliação Local (ajustado)	Kaggle (ajustado)
Árvore de Decisão	Ex: 0.7850	Ex: 0.7700	Ex: 0.8100	Ex: 0.7900
Naive Bayes				
KNN				

- Escreva de 3 a 5 linhas comentando os resultados

Entregar no SIGAA:

1. O notebook com o código, avaliação local e ajustes de hiperparâmetros;

2. Capturas de tela do Kaggle mostrando os scores alcançados;
3. Uma tabela comparando:
 - Score local (treino/teste) e score no Kaggle para cada modelo;
 - Score antes e depois do ajuste de hiperparâmetros;
4. Um breve comentário (3 a 5 linhas) analisando os resultados e justificando as diferenças observadas.