

CURSO SISTEMAS DE INFORMAÇÃO

MINERAÇÃO DE DADOS

ATIVIDADE

Acesse o site: <https://www.openml.org>

1. Escolha um dataset adequado para uma tarefa de classificação supervisionada.
2. Escolha individual:
 - Cada aluno deverá selecionar um dataset diferente dos demais colegas.
 - (Verifique com o professor se outro aluno já escolheu o mesmo conjunto de dados.)
3. Treinamento do modelo:
 - Carregue o dataset escolhido.
 - Realize a preparação dos dados (limpeza e divisão em treino e teste).
 - Treine um modelo de classificação de sua escolha (ex.: Árvore de Decisão, Regressão Logística, KNN, etc.).
4. Avaliação do modelo:
 - Após o treinamento, calcule e explique o significado de cada uma das métricas abaixo:
 - Matriz de Confusão
 - Acurácia (Accuracy)
 - Precisão (Precision)
 - Revocação (Recall)
 - F1-Score
 - Curva ROC e AUC
5. Entrega:
 - Apresente os resultados obtidos em um notebook e escreva uma breve interpretação sobre o desempenho do modelo.