

Trabalho Prático Final: Algoritmos Supervised Machine Learning para aprendizagem de modelos de classificação de dados

Valor: 6 valores (a teórica terá um valor total de 14 valores) | Mínimos: 30%

Data de entrega: 07-02-2026 | Data da defesa: 13-02-2026

Grupos: 3 elementos | Defesa: presencial com a participação de todo o grupo

Descrição: Desenvolva em Python, utilizando as bibliotecas numpy, pandas, sklearn e outras que considere relevantes, um conjunto de programas que permita determinar a exatidão (accuracy) e a matriz de confusão (confusion matrix) de cada um dos modelos machine learning criados utilizando os seguintes algoritmos:

- 1. Naive Bayes;
- 2. Decision Tree;
- 3. Random Forest;
- 4. kNN

Estes algoritmos de aprendizagem devem ser aplicados a dois datasets, disponíveis nos sites da [UCI](#) e do [kaggle](#) na construção de cada um dos modelos machine learning para classificação de dados:

- 1. [Breast Cancer](#)
- 2. [Heart Failure](#)

O programa deve garantir a preparação dos dados de forma a permitir o uso na construção de diferentes modelos por diferentes algoritmos. Verifique a evolução dos resultados quando utiliza diferentes percentagens do set de testes:

- 1. 15%
- 2. 30%
- 3. 50%

Baseie-se na tabela em baixo para apresentar um resumo dos resultados obtidos:

Precision	Naive Bayes	Decision Tree	Random Forest	kNN
labelencoder				
labelEncoder+standardScaler				
labelEncoder+oneHotEncoder				
labelEncoder+oneHotEncoder+standardScaler				

Deve ser entregue o código desenvolvido, bem como um relatório explicativo do trabalho realizado. Espera-se uma análise crítica dos resultados obtidos.