



**Universidade Federal da Bahia**

**Escola Politécnica**

**Departamento de Engenharia Elétrica e de Computação**

**Disciplina: Intro. ao Aprendizado de Máquina (ENGG67)**

**Professor: Antônio C. L. Fernandes Jr.**

**Aluno(a): \_\_\_\_\_**

**Data de apresentação e entrega: 05/11/2024**

**Horário limite para carregamento no AVA-MOODLE: 23:59**

## **Avaliação 01**

### Observações importantes:

1. O trabalho pode ser executado em equipe de duas pessoas.
2. O texto e o código devem estar em um arquivo de JUPYTER NOTEBOOK.
3. A entrega do trabalho deve ser realizada em um único arquivo do tipo “.ipynb”, carregado na área específica do AVA-Moodle da disciplina.
4. O nome do arquivo deve seguir o formato “IAM\_AVA01\_nomeestudentex\_nomeestudentey.ipynb”.

## **Questões**

1ª Questão: (10,0 pontos) Utilizando o banco de dados “Heart Disease”:

(<https://archive.ics.uci.edu/dataset/45/heart+disease>) faça o que se pede a seguir.

- a. O conjunto de dados é univariado ou multivariado? Descreva a quantidade de atributos preditivos e atributos alvo.
- b. Caracterize os dados quanto ao Tipo e Escala;
- c. Verifique a necessidade de balanceamento dos dados e caracterize estatisticamente os dados (calcule, para cada atributo, suas estatísticas univariadas pertinentes dentre frequência, moda, média, mediana, desvio-padrão, Q1, Q3, obliquidade e curtose).
- d. Analise os dados quanto à sua incompletude, inconsistências, redundâncias e presença de ruído;
- e. Calcule estatísticas multivariadas de covariância e correlação para os atributos quantitativos presentes. Interprete os valores obtidos.
- f. Trace boxplots dos atributos preditivos e identifique a presença ou não de outliers.
- g. Plote histogramas dos atributos e interprete sua distribuição, considerando as classes.
- h. Promova uma transformação de dados consistente com o banco sob análise;
- i. Utilize a análise de componentes principais (PCA) para gerar gráficos no espaço de atributos.