



Universidade Federal da Bahia

Escola Politécnica

Departamento de Engenharia Elétrica e de Computação

Disciplina: Introdução ao Aprendizado de Máquina (ENGG67)

Professor: Antônio C. L. Fernandes Jr.

Aluno(a): _____

Data: 02/05/2023

Avaliação 01

Observações importantes:

1. O trabalho deve ser executado em equipe de duas pessoas.

Questões

1ª Questão: (10,0 pontos) Utilizando o banco de dados “Iris Flower Data Set” faça o que se pede a seguir. Utilize todas as estratégias de visualização de dados (boxplots, gráficos de dispersão etc), todos os parâmetros estatísticos (média, desvio padrão, curtose etc) que você julgar pertinentes para cumprir as tarefas propostas abaixo.

- a. Caracterize os dados quanto ao Tipo e Escala;
- b. O conjunto de dados é univariado ou multivariado?
- c. Verifique a necessidade de balanceamento dos dados e caracterize estatisticamente os dados;
- d. Analise os dados quanto à sua incompletude, inconsistências, redundâncias e presença de ruído;
- e. Promova uma transformação de dados consistente com o banco sob análise;
- f. Utilize a análise de componentes principais (PCA) para gerar gráficos no espaço de atributos.