



FUNDAÇÃO EDUCACIONAL SERRA DOS ÓRGÃOS - FESO  
CENTRO UNIVERSITÁRIO SERRA DOS ÓRGÃOS - UNIFESO  
CENTRO DE CIÊNCIAS E TECNOLOGIA - CCT  
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO  
DISCIPLINA: SISTEMAS INTELIGENTES  
PROFESSOR: HERMANO L. S. LUSTOSA

## Trabalho

- O trabalho a seguir deve ser desenvolvido individualmente ou em grupo de no máximo 2 alunos.
- Os alunos devem utilizar um programa em uma linguagem de programação conhecida, que execute o algoritmo K-means com base no conjunto de dados fornecido.
- Os alunos deverão enviar o código fonte utilizado, junto com um documento com os resultados para o email: [hllustosa@gmail.com](mailto:hllustosa@gmail.com)

Neste trabalho, os alunos devem utilizar o algoritmo k-means sobre um conjunto de dados sobre flores e avaliar os resultados obtidos.

O conjunto de dados fornecidos tem informações sobre as medidas de algumas flores. Existem 4 dados básicos sobre cada flor, juntos com a espécie da planta. Os 4 atributos de dados fornecidos são:

- Largura da pétala
- Comprimento da pétala
- Largura da sépala
- Comprimento da sépala

Cada flor tem uma espécie, que pode ser:

- Iris-setosa
- Iris-versicolor
- Iris-virginica

A cada grupo será atribuído um par dessas variáveis, e este par deverá ser usado para a execução do algoritmo K-means (com  $k = 3$ ). Os alunos devem executar 3 experimentos considerando 3 configurações iniciais de centroides diferentes. As configurações devem ser escolhidas como os alunos preferirem. Ao final, devem ser gerados gráficos com os resultados fornecidos pelo algoritmo. Para cada experimento também é necessário contabilizar o porcentagem de ocorrências de cada uma das

especíes em cada grupo, e indicar em qual dos experimentos, obteve-se uma melhor classificação, isto é, qual configuração inicial forneceu um agrupamento mais próximo do esperado, colocando espécimes iguais no mesmo grupo.