

**Alunos:** Guilherme Diel e Luigi Belló Boscatto.

## **Tarefa 1B de TEG**

### **Avaliação de um grafo e cálculo de sua acurácia**

O trabalho consiste na implementação de um código que analisa as arestas geradas entre os vértices (flores) da Iris Dataset, quando a distância entre eles for menor ou igual a 0,3. Quando as arestas são formadas, as flores são divididas em dois grupos (setosa e não-setosa) de acordo com a existência ou não de aresta entre elas. Nessa análise, existem quatro casos possíveis para cada ligação:

- Verdadeiro Positivo: vértices (flores) são da mesma espécie e estão no mesmo grupo;
- Verdadeiro Negativo: vértices (flores) são de espécies diferentes e estão em grupos diferentes;
- Falso Positivo: vértices (flores) são de espécies diferentes, porém estão no mesmo grupo.
- Falso Negativo: vértices (flores) são da mesma espécie, porém estão em grupos diferentes;

Definidas as possibilidades, a função itera entre os todas as arestas do arquivo .txt e compara o primeiro vértice (se estiver entre as 50 primeiras flores) com o segundo vértice que forma a aresta. Se porventura o algoritmo encontrar um dos casos citados acima, o contador próprio do caso é aumentado.

Ao final da função, o algoritmo calcula uma das métricas de qualidade da classificação: a acurácia – através da fórmula  $(TP + TN) / (TP + FP + TN + FN)$ .

Observações importantes:

- Programa funciona no Linux, sendo buildado através do comando “gcc main.c -o a -lm”;
- Tentamos utilizar o DFS (Depth First Search) para separar os vértices em dois grupos, porém o resultado foi apenas um grande grupo com todos os vértices juntos. Dessa maneira, não era possível realizar qualquer tipo de análise, então separamos os grupos manualmente por setosas e não-setosas.