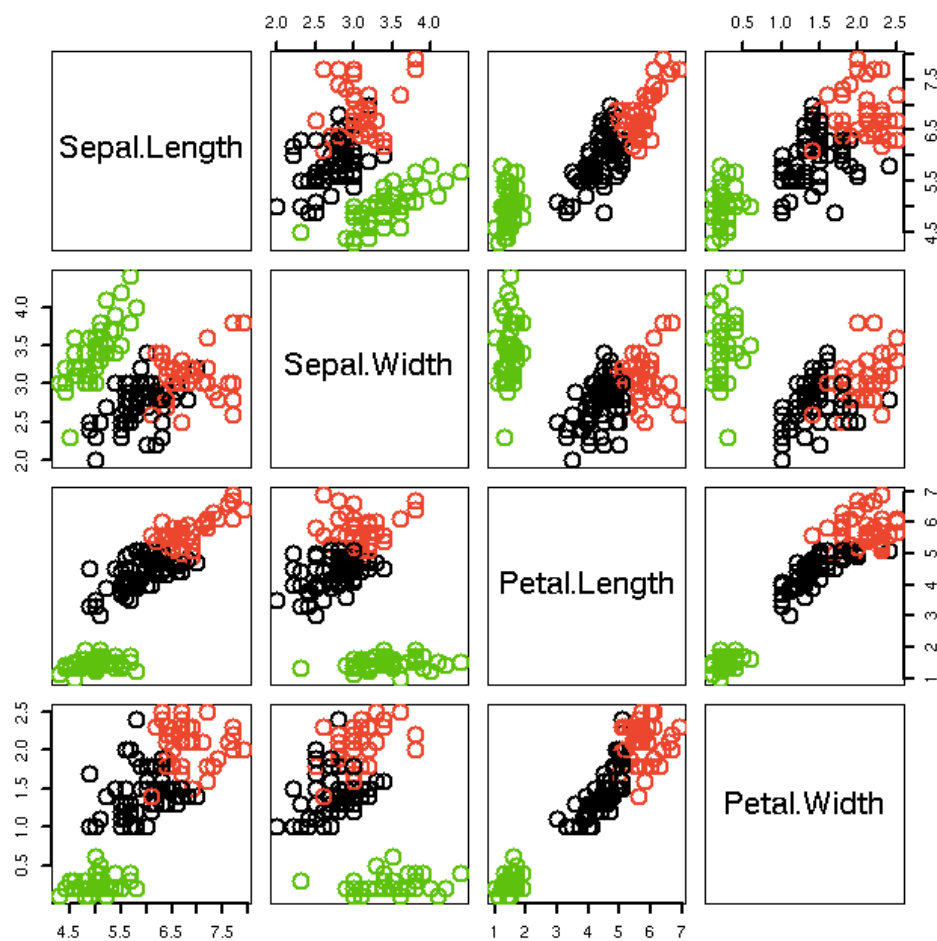


Universidade Federal de São Carlos  
Departamento de Computação  
Disciplina de Inteligência Artificial  
Prof. Ricardo Cerri

Segundo Trabalho Prático  
**Data de entrega: 14/12/2017**

**Tarefa:** escolha um dos conjuntos de dados abaixo para fazer um agrupamento utilizando o algoritmo k-means. Como os conjuntos de dados são para classificação, vocês devem verificar quantas classes possui o conjunto de dados escolhido. O número de grupos do agrupamento será então esse número de classes. Vejam um exemplo com o conjunto de dados Iris. Lembrem-se que o atributo alvo não participa do agrupamento.



Veja que, como o conjunto de dados Iris possui quatro atributos, o gráfico mostra as partições para todas as combinações de pares de atributos.

Apresentar como resposta, em forma de relatório:

1. Gráfico bidimensional com o agrupamento encontrado pelo algoritmo k-means. Para os conjuntos de dados com mais de 5 atributos, utilizar apenas os 5 primeiros para gerar o gráfico;
2. Descrição do conjunto de dados escolhido, com estatísticas básicas:
  - a. Número de exemplos
  - b. Número de atributos
  - c. Número de atributos numéricos
  - d. Número de atributos categóricos
  - e. Número de atributos com valores ausentes

**Obs1:** os conjuntos de dados podem conter valores ausentes. Substituir os valores numéricos pela média e os valores categóricos pela moda.

**Obs2:** entregar o relatório junto com o código utilizado para a realização do trabalho.

**Obs3:** O kmeans deve ser implementado, ou seja, não podem ser usados pacotes prontos para implementação do algoritmo. Para gráficos e tratamento dos dados, podem ser usadas funções prontas.

**Obs4 :** Devem ser formados grupos de 2 ou 3 alunos!

#### **Lista de conjuntos de dados:**

1. <https://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+signs+%28High+Quality%29>
2. <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>
3. <https://archive.ics.uci.edu/ml/datasets/banknote+authentication>
4. <https://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center>
5. <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>
6. <https://archive.ics.uci.edu/ml/datasets/Cardiotocography>
7. <https://archive.ics.uci.edu/ml/datasets/Daphnet+Freezing+of+Gait>
8. <https://archive.ics.uci.edu/ml/datasets/Ecoli>
9. <https://archive.ics.uci.edu/ml/datasets/Glass+Identification>