

1- Desenvolver uma rede MLP para o problema de classificação do IRIS dataset. A base de dados pode ser encontrada em <http://archive.ics.uci.edu/ml/>

O problema a ser resolvido é um problema de classificação entre três classes de flores (setosa, virginica e versicolor). Cada exemplo de flor apresenta 4 atributos. O conjunto total é composto de 150 exemplos de flores, divididas nas 3 classes.

Recomendações:

- Divida os dados entre conjunto de treinamento e de teste. (100 e 50, por exemplo);
- Normalize todos os atributos entre 0 e 1. Para isso, antes de apresentar os dados à rede calcule o valor máximo (max) e o valor mínimo (min) de cada atributo utilizando todo o conjunto de dados. De posse destes valores, para cada nova amostra  $x_i = [x_{i1} \ x_{i2} \ x_{i3} \ x_{i4}]$ , deve-se fazer  $x_{ij} = (x_{ij} - \min_j) / (\max_j - \min_j)$  para todos os atributos  $j$ .
- Utilize 3 neurônios na camada de saída. As saídas desejadas deve ser [1 0 0], [0 1 0] ou [0 0 1].

2- Implementar o algoritmo k-médias para o problema de clustering utilizando os dados do IRIS dataset. A base de dados pode ser encontrada em <http://archive.ics.uci.edu/ml/>

O problema a ser resolvido consiste em utilizar o k-médias para agrupar os dados em 3 classes.

Importante:

- Dado que este é um problema de aprendizado não supervisionado, as classes verdadeiras (rótulos) não serão utilizadas no algoritmo k-médias.
- Os rótulos deverão ser utilizados somente para verificar o resultado do k-médias.