

Exercício 1: Iris Data Set

Faça o download do conjunto de dados Iris do *UCI Machine Learning Repository*:

<http://archive.ics.uci.edu/ml/datasets/Iris>

Este conjunto de dados tem 150 itens, sendo 50 de cada uma de três espécies da flor *Iris*, mostrada na foto ao lado.



As três espécies são *I. setosa*, *I. virginica* e *I. versicolor*

Para cada item foram obtidos quatro atributos: comprimento e largura da sépala e da pétala.

Você deve treinar uma rede *Perceptron* para reconhecer as três diferentes classes.

Divida aleatoriamente os exemplos em subconjuntos de treinamento, validação e testes (por exemplo: 70%/15%/15%).

Teste diferentes pesos de inicialização e taxas de aprendizado para encontrar os melhores valores. Lembre-se de repetir o experimento diversas vezes com cada configuração.

Entregue o relatório documentando seus passos, não se esqueça de mostrar os parâmetros iniciais e finais das redes ajustadas, a taxa de aprendizado, as taxas de acerto/erro. Inclua também um gráfico do erro médio quadrático mostrando a convergência do algoritmo, e outros gráficos e/ou tabelas que julgar relevantes.

Mostre uma tabela com as matrizes de confusão para cada subconjunto separadamente: treinamento, validação e teste. O que você pode concluir desses dados?

Exercício 2: Wine Data Set

Faça o download do conjunto de dados Wine do *UCI Machine Learning Repository*:

<http://archive.ics.uci.edu/ml/datasets/Wine>

Identifique quais são os atributos e classes do problema e coloque-os em seu relatório.



Realize experimentos similares aos realizados com o conjunto de dados Iris.

Exercício 3: Escolha seu Conjunto de Dados e Algoritmos

Realize experimentos de classificação ou qualquer outra tarefa de aprendizado de máquina com redes neurais em um conjunto de dados de sua escolha, que esteja disponível publicamente.

Qualquer rede neural pode ser utilizada e comparações entre diferentes redes neurais são encorajadas.

Observações:

- 1) Em alguns casos, melhores resultados podem ser obtidos se os atributos forem normalizados previamente. Para isso obtenha a média e desvio padrão de cada atributo (coluna) e então, para cada elemento, subtraia a média e divida o resultado pelo desvio padrão. Este procedimento é conhecido como **zscore** e existem funções prontas para realiza-lo em diversas linguagens de programação.
- 2) Qualquer linguagem de programação pode ser utilizada. Porém o código-fonte deve ser devidamente comentado.
- 3) O relatório deve ser entregue em um arquivo no formato PDF, através de upload no *Google Classroom*. Os códigos-fontes também devem ser incluídos no *Google Classroom*. Não é necessário anexar códigos no PDF do relatório.