

# Trabalho 4

---

*Data de entrega: 12/06/2013*

*Número máximo de acadêmicos por trabalho: 1*

Dado o conjunto de treinamento contido na pasta do trabalho, apresente a taxa de acerto dos algoritmos de árvores de decisão, KNN e perceptron utilizando validação cruzada com 10 dobras. O conjunto de treinamento a ser usado (disponível na página) contém informações provenientes de diagnósticos feitos em pessoas com tumores malignos e benignos. O intuito com este trabalho é verificar qual dos 3 algoritmos acima apresenta o melhor desempenho (tomando suas taxas de acerto como base) para esse cenário: identificar se um tumor é maligno ou benigno com base em algumas informações sobre tal tumor.

Considere que o conjunto de treinamento já está aleatorizado e respeite a ordem apresentada das instâncias. Na árvore de decisão, assuma o valor 2 como positivo e 4 como negativo. No algoritmo KNN, utilize a distância euclidiana como métrica de distância entre as instâncias e  $K=10$ . No algoritmo perceptron, utilize peso inicial de 0.5 em todas as conexões, taxa de aprendizagem de 0.2 e treine por 200 épocas (associe a saída 1 da rede com a classe 2 e a saída 0 com a classe 4).

Como resultado do trabalho, apresente as matrizes de confusão de cada algoritmo juntamente com a respectiva taxa de acerto total e taxa de acerto por classe (maligno e benigno). As instâncias de cada grupo da validação cruzada devem ser escolhidos sequencialmente (da primeira à última instância). Por fim, conclua com base nas taxas de acerto qual foi o melhor algoritmo para o caso geral e qual foi o melhor algoritmo para cada classe.

## Observações:

- O trabalho deverá ser apresentado ao professor na data de entrega.
- Qualquer cópia indevida anula a nota integral do trabalho.
- As implementações devem seguir os algoritmos de aprendizagem vistos em sala.
- Não é permitido o uso de bibliotecas de inteligência artificial.
- O trabalho pode ser implementado em qualquer linguagem de programação de domínio do acadêmico.