

## 2º Trabalho

Curso: Engenharia de Computação  
Disciplina: Inteligência Computacional  
Prof. Jarbas Joaci de Mesquita Sá Junior  
Universidade Federal do Ceará – UFC/Sobral

**Entrega:** 21/05/2018 via e-mail para jarbas\_joaci@yahoo.com.br e david.borges@protonmail.com (enviar os arquivos) – (**Obs.** 1,0 ponto por dia de atraso na entrega).

1. Classificar a base de dados “database\_pap.dat”, que contém 917 amostras divididas em 2 classes.

- Cada linha da base corresponde a uma amostra representada por um vetor de 8 atributos.
- As primeiras 242 linhas pertencem à classe 0 e as 675 linhas restantes pertencem à classe 1.
- Usar o método dos mínimos quadrados (ver slides Dermatologia.pdf) para classificação.
- Aplicar o método *leave-one-out* para obter a taxa de classificação média das amostras. *Por exemplo, se há 5 amostras no total, 1 amostra é reservada para teste e as 4 restantes para treinamento. O processo é repetido 5 vezes, cada vez com uma amostra de teste diferente. A classificação média nesse caso é a média das 5 taxas de acertos obtidas.*

2. Classificar a mesma base de dados com o classificador **KNN (K=1)**. Aplicar o método *hold-out* para obter a taxa de classificação média das amostras. 70% das amostras deverão ser usadas para *treinamento* e as 30% restantes para *teste*. Deverão ser efetuadas 10 rodadas. Em cada rodada, os dados deverão ser permutados para que novos conjuntos de treinamento e teste sejam obtidos. A classificação média nesse caso é a média das 10 taxas de acertos obtidas.

---

**Obs.:** dois experimentos deverão ser realizados: no primeiro, os vetores de características permanecem inalterados; no segundo, os vetores de características são normalizados para ficarem com média nula e variância unitária.

Elaborar pequeno relatório sobre os resultados obtidos.

As implementações devem ser em C, Java, ou SCILAB/MATLAB.

**Boa Diversão!**