



UNIVERSIDADE FEDERAL DO PIAUÍ
CAMPUS SENADOR HELVÍDIO NUNES DE BARROS
DISCIPLINA: SISTEMAS INTELIGENTES
PROFESSORA: DEBORAH MAGALHÃES
MONITORES: DIEGO E EDSON

TRABALHO COMPUTACIONAL

1. Descrição do Trabalho

A base de dados utilizada neste trabalho se refere à dados de biópsia em pacientes com câncer de mama. Tais dados foram obtidos nos Hospitais da Universidade de Wisconsin, Madison, do Dr. William H. Wolberg. Ele avaliou biópsias de tumores mamários para 699 pacientes até 15 de julho de 1992, onde cada um dos nove atributos foi marcado em uma escala de 1 a 10. O diagnóstico do paciente é conhecido. O dataset conta com 699 linhas e 11 colunas.

Formato

O quadro de dados contém as seguintes colunas:

ID - número de código da amostra (não exclusivo).

V1 - espessura do aglomerado.

V2 - uniformidade do tamanho da célula.

V3 - uniformidade da forma celular.

V4 - adesão marginal.

V5 - tamanho de célula epitelial única.

V6 - núcleos nus (falta de 16 valores).

V7 - cromatina blanda.

V8 - normal nucleoli.

V9 - mitoses.

class - "benign" or "malignant".

A base de dados está disponível no link:

<https://raw.githubusercontent.com/vincentarelbundock/Rdatasets/master/csv/MAS S/biopsy.csv>

Acesse o link e salve os dados em um arquivo CSV.

1.1. O que deve ser feito?

- Treinamento de modelo (Cross Validation)
- Gráfico de Previsão x Observado
- Valor do RMSE

Atente que a abordagem utilizada (redes neurais, regressão linear) fica ao seu critério, podendo ser utilizada as técnicas abordadas nas práticas.

2. Avaliação

Este trabalho corresponde a terceira avaliação parcial da disciplina e deverá ser entregue no dia **07/12**. A nota do trabalho é **individual** e assumirá o valor de **1-10**.

Os seguintes critérios serão considerados na avaliação:

1. Atender ao que foi pedido na descrição deste documento;
2. Código está executando sem erros;
3. Organização do código.

Atenção: se identificada a cópia de código, a nota **zero** será atribuída aos envolvidos.