



INSTITUTO DE
INFORMÁTICA

Programa de Mestrado em Ciência da Computação
Algoritmos de Reconhecimento de Padrões (ARP)

Prof. Dr. Gustavo Teodoro Laureano
Prof. Dr. Anderson da Silva Soares

2017/1

LISTA DE EXERCÍCIOS - CLASSIFICADOR BAYESIANO

A base de dados Íris (<https://archive.ics.uci.edu/ml/datasets/Iris>) é formada por amostras de flores Íris e contém os seguintes atributos:

1. Column 1: sepal length in cm (SL)
2. Column 2: sepal width in cm (SW)
3. Column 3: petal length in cm (PL)
4. Column 4: petal width in cm (PW)
5. Column 5: class: (C)
 - a. Iris Setosa (C1)
 - b. Iris Versicolour (C2)
 - c. Iris Virginica (C3)

Cada classe possui 50 amostras, contabilizando 150 instâncias. Para os problemas de classificação abaixo, considere 70% das amostras (35 primeiras amostras) de cada classe como dados de treinamento e 30% como dados de teste (15 últimas amostras).

Problema 01: Apresente o gráfico das funções de probabilidade: $p(\text{SL} | C_i)$, $p(\text{SW} | C_i)$, $p(\text{PL} | C_i)$, $p(\text{PW} | C_i)$ e o gráfico do modelo de classificação $p(C_i | \text{SL}, \text{SW}, \text{PL}, \text{PW})$ usando o classificador Bayesiano. **Nota:** as variáveis em questão são variáveis contínuas.

Problema 02: Realize a classificação dos dados de teste usando o classificador do Problema 02. Calcule e apresente a taxa de acerto de cada amostra de teste para cada classe.