

Tarefa 2 - data de entrega: até 10/09

Em situações em que existe conhecimento sobre a distribuição dos dados é possível utilizar estimadores de *Máxima Verossimilhança* para estimar os parâmetros dos modelos. Quando a distribuição é desconhecida, uma alternativa é testar uma série de distribuições e avaliar qual caracteriza melhor os dados, segundo algum critério. Nesta tarefa aplicaremos este procedimento para classificação.

## 1 Tarefa

A Tarefa 2 consiste em construir um classificador que diferenciará 3 classes  $C_1$ ,  $C_2$  e  $C_3$  utilizando um único atributo. Para isto, serão fornecidos uma amostra de 60 exemplos classificados (20 de cada classe) e um conjunto de validação contendo 60 exemplos classificados. Suspeita-se que as distribuições do atributo *em cada classe* sejam dados por distribuições normais, exponenciais ou uniformes, porém não se sabe qual das classes possui qual distribuição. Deve-se estimar os parâmetros de cada distribuição candidata para cada uma das classes utilizando *estimadores de Máxima Verossimilhança* e utilizar o *Classificador de Bayes* para a escolha dos modelos.

O desempenho de cada classificador será medido utilizando o número de erros cometidos no conjunto de testes.

### 1.1 Conjuntos de dados

Os dados de treinamento e teste estão disponíveis no PACA, juntamente com um script gerador de conjuntos de teste.

## 2 Entrega

A Tarefa deverá ser feita em Python. O pacote de computação científica SciPy possui funções estatísticas que podem ser úteis. Veja a documentação neste link e um pequeno tutorial sobre arrays. O SciPy é o principal pacote científico implementado em Python e possui API similar ao Matlab.

Além do programa, deverá ser feito um relatório que contenha os resultados de cada teste classes×distribuições e que aponte qual foi a melhor combinação. Como as distribuições possuem interseção, seria interessante calcular a matriz de confusão da combinação escolhida.

Tanto as implementações como o relatório devem ser entregues no PACA.

Dúvidas podem ser postadas no fórum de discussões da disciplina ou esclarecidas diretamente com o monitor Igor ou com a Profa.