

Universidade Federal do Ceará - Campus de Quixadá

## Prática 02

### BCC0019 - Aprendizado de Máquina

Prof. Carlos Igor Ramos Bandeira

Outubro de 2016

Aluno (Nome/Matrícula): \_\_\_\_\_

A PONTUAÇÃO TOTAL QUE PODE SER CONQUISTADA COM ESTE TRABALHO É DE 10,0.

PRAZO DE ENTREGA: DEZEMBRO/2016.

1. (80 points) Usando o conjunto de dados “coluna.dat”, implementar os seguintes classificadores:

- Distância euclidiana mínima (DEM):

$$g_i(x) = \|\mathbf{x} - \mu_i\|^2, \quad i = 1, 2, 3. \quad (1)$$

- Distância mahalanobis mínima (DMM):

$$g_i(x) = (\mathbf{x} - \mu_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mu_i), \quad i = 1, 2, 3. \quad (2)$$

em que  $\mu_i$  e  $\mathbf{C}^{-1}$  são, respectivamente, o centróide e a matriz de covariância da  $i$ -ésima classe.

Para cada classificador montar os resultados de acordo com o procedimento abaixo:

- Após 50 rodadas de treinamento e teste, para cada classificador calcular suas taxas de acerto (i) média, (ii) mínima, (iii) máxima e (iv) taxas de acerto médias por classe.
- Montar a matriz de confusão para cada classificador usando dados de testes (partição de 80% dos dados para treinar e 20% para testar).

Obs.1: Normalizar os atributos antes de utilizar o classificador DEM. Não é preciso normalizar os atributos para o classificador DMM.

Obs.2: Os parâmetros dos classificadores DEM e DMM devem ser estimados usando os dados de treinamento.

2. (20 points) Quais são as suposições teóricas feitas acerca dos dados quando se utiliza o classificador DEM? E para o classificador DMM?