

Universidade Federal do Ceará  
Campus Sobral  
Engenharia da Computação e Engenharia Elétrica

Tópicos Especiais em Telecomunicações I (ECO0080)  
(Reconhecimento de Padrões)

**Classificadores KNN e MDC com k-fold**

**0) Informações Gerais**

- Trabalho Individual. Apenas simulação, sem trabalho escrito.
- O código deve estar bem organizado e comentado, para que possa estar inteligível.
- Não usar “funções prontas” nem para os classificadores nem para o k-fold.
- Enviar o código / implementação (Matlab / Python...) para o email:

[david.coelho@sobral.ufc.br](mailto:david.coelho@sobral.ufc.br)

- Prazo para entrega: 18/06/2021 às 23:59

**1) Banco de dados: Sinais no tempo.**

- É composto por duas classes: Sinal de ECG (classe1) e Sinal de áudio (classe2).
- Cada classe possui 50 sinais (amostras), e cada sinal possui “duração de 500 pontos”.
- a) Gerar vetor de atributos a partir dos vetores de sinais.
- Usar pelo menos 5 atributos (Obs: verifiquem slides 29 e 30 da aula 02).
- Escolher os atributos que achar mais convenientes.

**2) Classificadores**

- a) Implementar os classificadores:

- | KNN | NPC (Nearest Prototype Classifier) = MDC (Minimum Distance to Centroid) |

### **3) Experimento com o algoritmo MDC (fazer 10 realizações)**

#### **a) Realização:**

- Dividir dados, aleatoriamente, entre treinamento e teste.

(Ex: 80% para treino, 20% para teste)

- Constrói o modelo com os dados de treinamento
- Verifica a acurácia (taxa de acerto) do modelo (teste).

#### **b) Estatística:**

- Calcular a taxa de acerto média das 10 realizações.

### **4) Experimento com o algoritmo KNN (fazer uma realização para cada valor de K)**

- Valores de  $K = 2$  à  $10$ .

#### **a) Realização**

- Para cada valor de  $K$ , utilizar a validação cruzada k-fold (5-fold):
  - Dividir, aleatoriamente, o banco de dados em 5 grupos
  - Repetir o seguinte processo 5 vezes: Constrói o modelo com 4 grupos, e testa com 1.  
(utilizando cada grupo, uma única vez, como conjunto de teste)
- Calcula a taxa de acerto médio dos procedimentos.

#### **b) Defina o melhor valor de $K$ (KNN) para o banco de dados**

- Fazer um comentário no código.