



UNIVERSIDADE FEDERAL DO CEARA
CAMPUS DE SOBRAL
ENGENHARIA DA COMPUTAÇÃO E ENGENHARIA ELÉTRICA

TÓPICOS ESPECIAIS EM TELECOMUNICAÇÕES I (ECO0080)

TEMA: RECONHECIMENTO DE PADRÕES

Prof.: C. Alexandre Rolim Fernandes

Trabalho 3 – Classificação com Árvore de Decisão

- Trabalho Individual
- Sugere-se o uso do MATLAB ou Python para a prática de simulação .
- O código deve estar bem organizado e comentado, para que seja possível entendê-lo e corrigi-lo. Códigos que estejam desorganizados ou sem os devidos comentários explicativos terão penalização na nota.
- Fazer todas as questões em **um só arquivo**.
- O seu código deve **gerar automaticamente todos os gráficos e resultados** solicitados.
- Enviar no SIGAA **apenas o arquivo do código**.
- **O nome do arquivo deve ser igual ao seu nome**.
- **Não enviar código em formato PDF ou de texto (DOC, ODT, TXT etc)**.
- **Não enviar arquivo compactado** (zip,rar etc).
- Prazo e forma de entrega: dia 16/08/24 às 23h59, no SIGAA.

Prática de Simulação - Classificação de uma Base de Dados com Atributos Categóricos:

- Implementar uma Árvore de Decisão para classificar a base de dados "Dataset" fornecida. Esta base de dados foi fornecida nos formatos mat (Matlab data) e csv (comma-separated values).

- Informações da base de dados:

- 876 amostras
- 3 atributos de entrada categóricos (binários):
 - empregado?: “sim” e “não”
 - devedor?: “sim” e “não”
 - salário acima de 5SM?: “sim” e “não”
- 2 classes de saída
 - empréstimo?: “sim” e “não”
 - 462 amostras da classe “não”
 - 414 amostras da classe “sim”
- “sim” corresponde ao 1 e “não” corresponde ao 0

- No algoritmo da Árvore de Decisão, você deve usar a entropia (ganho de informação) como critério de escolha dos nós.

- Uma vez que um atributo foi usado uma vez, não usar novamente este atributo em outros nós.

- O algoritmo deve se parar quando não for possível mais reduzir a entropia ou quando todos os 3 atributos forem usados.

- Não é necessário realizar podas.

- Usar validação cruzada K-fold com K=10

- Não há a etapa de extração de atributos. As entradas fornecidas já correspondem aos atributos.

- Não usar funções prontas do MATLAB para a árvore de decisão nem para o cálculo da entropia.

- O algoritmo deve ter como saída os seguintes parâmetros (média dos 10 folds):

- acurácia
- sensibilidade
- especificidade
- precisão
- f1-score