

Exercícios de Árvores de Decisão

1 – Dada a tabela abaixo com os Ganhos de Informação por nível, construa a árvore de decisão.

	X1	X2	X3	X4	X5
Primeiro Nível	0.030	0.022	0.031	0.040	0.015
Segundo Nível E	0.020	0.035	0.012	0.001	0.030
Segundo Nível D	0.043	0.032	0.050	0.032	0.045

Valores:

X1 = {M (D), F (D), I (E)}

X2 = {Alto (E), Baixo (D)}

X3 = {Sim (D), Não (E)}

X4 = 80..250 (Escolhido 135.2 como ponto de corte. D (\geq), E ($<$))

X5 = 1..100 (Escolhido 44.5 como ponto de corte. D ($<$), E (\geq))

D = Direita

E = Esquerda

Quando a árvore tiver montada, complete aleatoriamente as folhas com os valores $y \in \{0,1\}$

2 – Construa as Regras de Decisão da questão anterior.

3 – Dado o conjunto de treinamento abaixo, responda as perguntas:

a_1	a_2	classe
T	T	+
T	T	+
T	F	-
F	F	+
F	T	-
F	T	-

a) Qual é a entropia do conjunto de treinamento?

b) Qual é o ganho de informação do atributo a_2 ?

4 [CÓDIGO] – Acessar o *UCI Repository* e baixar duas bases de classificação. Treinar uma árvore de decisão para cada base com 80% dos dados. Os outros 20% usar para testes. Calcular a taxa de acerto do conjunto de testes de cada base de testes. Usar como modelo o notebook: *2_Árvores_de_Decisão.ipynb*

Link do *UCI Repository*: <https://archive.ics.uci.edu/>