

Agrupamento - Prática 3

Lorena da Silva Elias 11721BCC019
Victor Hugo Eustácio Lopes 11721BCC011
Yan Stivaletti e Souza 11821BCC002
Yago Vinícius Ferreira de Castro 11721BCC020

1. **Faça a binarização, usando codificação 1-de-n, dos seguintes valores:**
Amarelo, vermelho, verde, azul, laranja, branco

Valor categórico	x1	x2	x3	x4	x5	x6
amarelo	1	0	0	0	0	0
vermelho	0	1	0	0	0	0
verde	0	0	1	0	0	0
azul	0	0	0	1	0	0
laranja	0	0	0	0	1	0
branco	0	0	0	0	0	1

2. **Discretizar em 3 intervalos o atributo que possui os valores a seguir:**
0, 1, 3, 6, 6, 9, 10, 10, 10, 13, 18, 20, 21, 21, 25

Usar:

a) Larguras iguais

Largura = $(25 - 0) / 3 = 8,333...$

Para facilitar a visualização e incluir todos os valores usamos Largura = 9

- Bin 1: 0, 1, 3, 6, 6, 9 [0, 10[
- Bin 2: 10, 10, 10, 13, 18 [10, 20[
- Bin 3: 20, 21, 21, 25 [20, 30[

b) Frequências iguais

- Bin 1: 0, 1, 3, 6, 6
- Bin 2: 9, 10, 10, 10, 13
- Bin 3: 18, 20, 21, 21, 25

3. Escreva um pseudocódigo para representar a técnica de amostragem aleatória sem reposição.

Considere um dataset D de tamanho N.

Considere que se quer uma amostra A de tamanho M do dataset D.

1. $i \leftarrow 0$
2. Para D_i gere um número aleatório $R \in [0, 1]$
3. Se $R < M/N$ então
 - 4.1. $A \leftarrow A \cup \{D_i\}$
 - 4.2. $N \leftarrow N - 1$
 - 4.3. $M \leftarrow M - 1$Senão
 - 4.1. $N \leftarrow N - 1$
4. $i \leftarrow i + 1$
5. Se $N = 0$ ou $M = 0$ pare.
Senão retorne ao passo 2.