

# LISTA 6- aluno: lasmin Oliveira.

1)

Cálculo do ganho de informação (ID3).

Entropia total  $H(S) = 0.997503$

Atributo: **Experiência**

Valor	Nº exemplos	Gosta	Não Gosta	H(partição)	peso*H
Baixa	6	1	5	0.650022	0.229420
Média	6	4	2	0.918296	0.324104
Alta	5	4	1	0.721928	0.212332

Ganho(Experiência) =  $H(S) - \sum(p_i * H_i) = 0.231647$

Atributo: **Interesse**

Valor	Nº exemplos	Gosta	Não Gosta	H(partição)	peso*H
Baixo	10	3	7	0.881291	0.518406
Alto	7	6	1	0.591673	0.243630

Ganho(Interesse) =  $H(S) - \sum(p_i * H_i) = 0.235466$

Atributo: **Horas**

Valor	Nº exemplos	Gosta	Não Gosta	H(partição)	peso*H
Baixas	9	4	5	0.991076	0.524687
Altas	8	5	3	0.954434	0.449145

Ganho(Horas) =  $H(S) - \sum(p_i * H_i) = 0.023670$

**Resposta:** Raiz = Interesse; ganho = 0.235.

2)

Naive Bayes sem suavização para (Experiência=Alta, Interesse=Alto, Horas=Baixas).

Termo	Valor
P(Gosta)	$9/17 = 0.529412$
P(Não Gosta)	$8/17 = 0.470588$
P(Alta Gosta)	$4/9 = 0.444444$
P(Alta Não Gosta)	$1/8 = 0.125000$
P(Alto Gosta)	$6/9 = 0.666667$
P(Alto Não Gosta)	$1/8 = 0.125000$
P(Baixas Gosta)	$4/9 = 0.444444$
P(Baixas Não Gosta)	$5/8 = 0.625000$

Verossimilhança proporcional:  $P(\text{Gosta}) * \prod P(x_i | \text{Gosta}) = 0.069717$ ;  $P(\text{Não Gosta}) * \prod P(x_i | \text{Não Gosta}) = 0.004596$ .

Posteriores normalizados:  $P(\text{Gosta} | x) = 93.82\%$  ;  $P(\text{Não Gosta} | x) = 6.18\%$ .

**Resposta:** 93,82% e 6,18%.

**3)**

Questão experimental com execução de modelos no conjunto Titanic. A resposta depende da implementação/execução. (Não há contas a apresentar aqui.)

**4)**

Os cálculos de Apriori (Itemsets e Regras) requerem a tabela de transações citada na questão. Ela não está visível no PDF fornecido. Assim que as transações forem fornecidas, calculo suporte/confiança e respondo.