

QUESTÃO 1

1. Entropia Total (Ponto de Partida)

- Contagens (Gosta, Não Gosta): **(9, 8)**
- Total de exemplos: **17**

$$H(S) = -[(9/17) \cdot \log_2(9/17) + (8/17) \cdot \log_2(8/17)] = 0,9975$$
$$H(S) = -[(9/17) \cdot \log_2(9/17) + (8/17) \cdot \log_2(8/17)] = 0,9975$$

2. Avaliação dos Atributos e Ganhos

A. Atributo: Experiência

- Entropia pós-split (ponderada):
 $(6/17) \cdot 0,6500 + (6/17) \cdot 0,9183 + (5/17) \cdot 0,7219 = 0,7659$
 $(6/17) \cdot 0,6500 + (6/17) \cdot 0,9183 + (5/17) \cdot 0,7219 = 0,7659$
- **Ganho(Experiência):**
 $0,9975 - 0,7659 = 0,2316$

B. Atributo: Interesse

- Entropia pós-split (ponderada):
 $(10/17) \cdot 0,8813 + (7/17) \cdot 0,5917 = 0,7620$
 $(10/17) \cdot 0,8813 + (7/17) \cdot 0,5917 = 0,7620$
- **Ganho(Interesse):**
 $0,9975 - 0,7620 = 0,2355$

C. Atributo: Horas

- Entropia pós-split (ponderada):
 $(10/17) \cdot 0,9710 + (7/17) \cdot 0,8631 = 0,9266$
 $(10/17) \cdot 0,9710 + (7/17) \cdot 0,8631 = 0,9266$
- **Ganho(Horas):**
 $0,9975 - 0,9266 = 0,0709$

3. Conclusão

O atributo com **maior Ganho de Informação** será utilizado como **nó raiz** da árvore de decisão.

- **Melhor Atributo:** Interesse
- **Ganho Máximo:** 0,2355 (aprox. 0,235) -> **Alternativa correta: (b)**

QUESTÃO 2

1. Priors

- $P(\text{Gosta}) = 9/17$
- $P(\text{Não Gosta}) = 8/17$

2. Likelihoods

- Experiência:
 - $P(\text{Exp} = \text{Alta} \mid \text{Gosta}) = 4/9$
 - $P(\text{Exp} = \text{Alta} \mid \text{Não}) = 1/8$
- Interesse:
 - $P(\text{Int} = \text{Alto} \mid \text{Gosta}) = 6/9$
 - $P(\text{Int} = \text{Alto} \mid \text{Não}) = 1/8$
- Horas:
 - $P(\text{Horas} = \text{Baixas} \mid \text{Gosta}) = 4/9$
 - $P(\text{Horas} = \text{Baixas} \mid \text{Não}) = 5/8$

3. Scores (proporcionais à posterior)

- Gosta: $(9/17) * (4/9) * (6/9) * (4/9) = 32/459 \approx 0,0697168$

- Não: $(8/17) * (1/8) * (1/8) * (5/8) = 5/1088 \approx 0,0045956$

4. Normalização

- $P(\text{Gosta} | x) = (32/459) \div (32/459 + 5/1088) = 2048/2183 \approx 93,82\%$
- $P(\text{Não} | x) = (5/1088) \div (32/459 + 5/1088) = 135/2183 \approx 6,18\%$

5. Conclusão

- Alternativa correta: a) 93,82% e 6,18%

QUESTÃO 3

- O Random Forest apresentou um ganho significativo de desempenho após o ajuste de hiperparâmetros, especialmente para a classe 1 (sobreviventes), que geralmente é a mais desafiadora.
- A Decision Tree também teve boa melhora com o tuning e se mantém uma opção rápida de execução.
- O GaussianNB é extremamente veloz, porém suas métricas de performance ficaram abaixo das demais abordagens.

QUESTÃO 4

Itemsets de tamanho 1 frequentes: 3

(Café — sup=3, Pão — sup=5, Manteiga — sup=5)

Itemsets de tamanho 2 frequentes: 3

(Café, Manteiga — sup=3; Café, Pão — sup=3; Manteiga, Pão — sup=4)

Itemsets de tamanho 3 frequentes: 1

(Café, Manteiga, Pão — sup=3)

Número de regras com confiança ≥ 0.8 (apenas a partir dos itemsets frequentes): 7

Regras (suporte, antecedente \rightarrow consequente, confiança)

(Café) \rightarrow (Manteiga), sup=3, conf = $3/3 = 1.0$

(Café) \rightarrow (Pão), sup=3, conf = $3/3 = 1.0$

(Manteiga) \rightarrow (Pão), sup=4, conf = $4/5 = 0.8$

(Pão) \rightarrow (Manteiga), sup=4, conf = $4/5 = 0.8$

(Café) \rightarrow (Manteiga, Pão), sup=3, conf = $3/3 = 1.0$

(Café, Manteiga) \rightarrow (Pão), sup=3, conf = $3/3 = 1.0$

(Café, Pão) \rightarrow (Manteiga), sup=3, conf = $3/3 = 1.0$

QUESTÃO 5

Código

RESULTADO:

ITEMSETS FREQUENTES (min_sup = @.3) ===

('Pao",) | sup=@.608 | count=6

('café",) | sup=2.5@ | count=5

('Manteiga',) | sup=@.50@ | count=5

('Arroz',) | sup=@.3@8 | count=3

('Café', 'Manteiga') | sup=@.4e2 | count=4

('Café", P30") | sup=@.400 | count=4

('Manteiga', 'P3o") | sup=@.400 | count=4

(cae"

» 'Manteiga', 'Pao') | sup=-@.3e@ | count=3

=== REGRAS (min_conf = @.6) ===

Café => Po | sup=@.400 conf=9. lift=1.333

Café => Manteiga | sup=0.400 conf=0.800 lift=1.600

Manteiga => Café | sup=@.400 conf=@.800 lift=1.600

Manteiga => Po | sup=@.40@ conf=0.800 lift-1.333

Café, Manteiga => Pao | sup=0.30@ conf=0.75@ lift=1.250

Café, P3o => Manteiga | sup=0.300 conf=8.750 lift=1.500

Manteiga, Pao => Café | sup-0.300 conf=0.75@ lift-1.500

Po => Café | sup-0.400 conf-0.667 lift-1.333

Po => Manteiga | sup-0.400 conf-0.667 lift-1.333

Café => Manteiga, Pao | sup-0.300 conf=9.600 lift-1.500

Manteiga => Café, Pao | sup=0.300 conf=9.600 lift=1.500

QUESTÃO 6

Código

RESULTADO:

| sup0.600 (count=6)
(café",) | sup-@.580 (count=5)
(Manteiga',) | sup-@.58@ (count=5)
(CArroz',) | sup-@.380 _ (count=3)
(CCerveja',) | sup=6.200 (count=2)
(‘Feijao",) | sup=@.200 (count=2)
(CLeite',) | sup-0.200 (count=2)

Lt (Frequentes)

("P30",) | sup=@.600 (count=6)
(café",) | sup=0.580 (count=5)
(Manteiga',) | sup-@.58@ (count=5)
(Arroz',) | sup-8.388 (count=3)

2 (candidatos)

(‘café", “Manteiga’) | sup-0.40@ (count~4)
("Café", “P50") | sup-@.42 (count=4)
(Manteiga', "P30") | sup-@.48@ (count=4)
(Arroz', ‘P30") | sup-@.100 (count=1)
(CArroz', *Café") | sup=0.00@ (count=8)

(CArroz', ‘Manteiga’) | sup=0.000 (count=0)

L2 (frequentes)

café", 'Manteiga') | sup=.40e (count=4)

('café", *P30") | sup=0.40@ (count=4)

(CManteiga', 'P3o') | sup=0.400 (count=4)

3 (candidatos)

('Café", 'Manteiga', "Pao") | sup-@.300 (count=3)

L3 (Frequentes)

(Café", 'Manteiga', 'P3o') | sup=0.300 (count=3)

RESUMD: TODOS OS ITEMSETS FREQUENTES

("P50",) | sup=@.600

(Ccafé",) | sup-2.500

(CManteiga',) | sup=2.500

(arroz",) | sup:

('Café*', 'Manteiga') | sup-8.400

((café", *P30") | sup=2.400

(Manteiga', "Pa0") | sup-2.400

(/café", 'Manteiga', 'P50") | sup-0.300

QUESTÃO 7

Código

RESULTADO:

L4 (frequentes)

{não Arroz, Café, não Cerveja, Pão} | sup=0.400 count=4

{não Arroz, Café, não Feijão, Manteiga, Pão} | sup=0.400 count=4

{não Arroz, não Feijão, Manteiga, Pão} | sup=0.400 count=4

{não Café, não Cerveja, não Leite, não Manteiga} | sup=0.400 count=4

{Arroz, não Café, não Cerveja, não Leite} | sup=0.300 count=3

{não Arroz, não Café, não Cerveja, Manteiga} | sup=0.300 count=3

{Arroz, não Leite, não Manteiga, Pão} | sup=0.300 count=3

{não Arroz, não Cerveja, não Leite, não Manteiga} | sup=0.300 count=3

{não Cerveja, Manteiga, Pão} | sup=0.300 count=3

{Café, não Cerveja, não Feijão, Manteiga} | sup=0.300 count=3

{não Café, não Cerveja, não Feijão, Pão} | sup=0.300 count=3

{Café, não Cerveja, não Leite, Pão} | sup=0.300 count=3

{Café, não Feijão, Manteiga, Pão} | sup=0.300 count=3

{Café, não Feijão, não Leite, Manteiga} | sup=0.300 count=3

{não Arroz, Café, Manteiga, Pão} | sup=0.300 count=3

{não Arroz, Café, não Cerveja, Manteiga} | sup=0.300 count=3

{não Arroz, Café, não Leite, não Feijão} | sup=0.300 count=3

{não Arroz, Café, não Cerveja, não Leite} | sup=0.300 count=3

{não Arroz, Café, não Cerveja, não Feijão, Pão} | sup=0.300 count=3

{não Arroz, Café, não Feijão, Leite} | sup=0.300 count=3

{não Arroz, Café, não Leite, Manteiga} | sup=0.300 count=3

{não Arroz, Café, não Leite, Pão} | sup=0.300 count=3

{não Arroz, não Cerveja, Manteiga, Pão} | sup=0.300 count=3

{não Arroz, não Cerveja, não Feijão, Manteiga} | sup=0.300 count=3

{não Arroz, não Cerveja, não Feijão, Pão} | sup=0.300 count=3

{não Arroz, não Cerveja, não Feijão, não Leite} | sup=0.300 count=3

{não Arroz, não Cerveja, não Leite, Pão} | sup=0.300 count=3

{não Arroz, não Feijão, não Leite, Manteiga} | sup=0.300 count=3

{não Café, não Cerveja, não Feijão, não Leite} | sup=0.300 count=3

{não Café, não Cerveja, não Feijão, não Manteiga} | sup=0.300 count=3

{não Café, não Cerveja, não Leite, não Pão} | sup=0.300 count=3

{não Café, não Cerveja, Manteiga, não Pão} | sup=0.300 count=3

{não Café, não Feijão, não Leite, Manteiga} | sup=0.300 count=3

{não Café, não Leite, não Manteiga, não Pão} | sup=0.300 count=3

{não Cerveja, não Feijão, Manteiga, Pão} | sup=0.300 count=3

{não Cerveja, não Feijão, não Leite, Pão} | sup=0.300 count=3

{não Cerveja, não Feijão, não Leite, não Manteiga} | sup=0.300 count=3