

Lista 6 de IA - lasmin Oliveira

1) ID3 – Raiz e ganho do atributo:

Conjunto S (17 amostras): Gosta=9, Não Gosta=8.

Entropia(S) = $-(9/17) \cdot \log_2(9/17) - (8/17) \cdot \log_2(8/17) = 0.997503$ bits.

Experiência:

Alta: $|S_v|=5 \rightarrow$ (Gosta=4, Não Gosta=1) Entropia=0.721928; peso=5/17=0.294118; peso·entropia=0.212332.

Baixa: $|S_v|=6 \rightarrow$ (Gosta=1, Não Gosta=5) Entropia=0.650022; peso=6/17=0.352941; peso·entropia=0.229420.

Média: $|S_v|=6 \rightarrow$ (Gosta=4, Não Gosta=2) Entropia=0.918296; peso=6/17=0.352941; peso·entropia=0.324104.

Entropia condicional = 0.765856 \Rightarrow Gain(S,Experiência) = $0.997503 - 0.765856 = 0.231647$.

Interesse:

Alto: $|S_v|=7 \rightarrow$ (Gosta=6, Não Gosta=1) Entropia=0.591673; peso=7/17=0.411765; peso·entropia=0.243630.

Baixo: $|S_v|=10 \rightarrow$ (Gosta=3, Não Gosta=7) Entropia=0.881291; peso=10/17=0.588235; peso·entropia=0.518406.

Entropia condicional = 0.762036 \Rightarrow Gain(S,Interesse) = $0.997503 - 0.762036 = 0.235466$.

Horas:

Altas: $|S_v|=8 \rightarrow$ (Gosta=5, Não Gosta=3) Entropia=0.954434; peso=8/17=0.470588; peso·entropia=0.449145.

Baixas: $|S_v|=9 \rightarrow$ (Gosta=4, Não Gosta=5) Entropia=0.991076; peso=9/17=0.529412; peso·entropia=0.524687.

Entropia condicional = 0.973833 \Rightarrow Gain(S,Horas) = $0.997503 - 0.973833 = 0.023670$.

\Rightarrow A maior informação é para Interesse (ganho = 0.235). Raiz = Interesse.

2) Naive Bayes – Probabilidades para (Experiência=Alta, Interesse=Alto, Horas=Baixas):

$P(\text{Gosta}) = 9/17 = 0.529412$; $P(\text{Não Gosta}) = 8/17 = 0.470588$.

Verossimilhanças (sem Laplace):

$P(\text{Alta}|\text{Gosta}) = 4/9$, $P(\text{Alto}|\text{Gosta}) = 6/9$, $P(\text{Baixas}|\text{Gosta}) = 4/9$.

$P(\text{Alta}|\text{Não Gosta}) = 1/8$, $P(\text{Alto}|\text{Não Gosta}) = 1/8$, $P(\text{Baixas}|\text{Não Gosta}) = 5/8$.

$\text{Score_Gosta} = (9/17) * (4/9) * (6/9) * (4/9) = 0.06971678$

$\text{Score_NãoGosta} = (8/17) * (1/8) * (1/8) * (5/8) = 0.00459559$

Normalização Z = $0.06971678 + 0.00459559 = 0.07431236$

$P(\text{Gosta}|x) = \text{Score_Gosta} / Z = 0.06971678 / 0.07431236 = 93.815850\%$

$P(\text{Não Gosta}|x) = \text{Score_NãoGosta} / Z = 0.00459559 / 0.07431236 = 6.184150\%$

4) Apriori – suporte mínimo 0,3 e confiança 0,8.

Transações (10): conforme tabela (Sim = item presente).

Itemsets frequentes:

- 1-itemsets (suporte): [('Café', 0.3), ('Pão', 0.5), ('Manteiga', 0.5)] → total = 3
- 2-itemsets: [('Café, Pão', 0.3), ('Café, Manteiga', 0.3), ('Manteiga, Pão', 0.4)] → total = 3
- 3-itemsets: [('Café, Manteiga, Pão', 0.3)] → total = 1

Regras com confiança $\geq 0,8$ (conf, sup):

$\{\text{'Café'}\} \Rightarrow \{\text{'Pão'}\}$ (conf=1.00, sup=0.30)

$\{\text{'Café'}\} \Rightarrow \{\text{'Manteiga'}\}$ (conf=1.00, sup=0.30)

$\{\text{'Manteiga'}\} \Rightarrow \{\text{'Pão'}\}$ (conf=0.80, sup=0.40)

$\{\text{'Pão'}\} \Rightarrow \{\text{'Manteiga'}\}$ (conf=0.80, sup=0.40)

$\{\text{'Café'}\} \Rightarrow \{\text{'Pão'}, \text{'Manteiga'}\}$ (conf=1.00, sup=0.30)

$\{\text{'Café'}, \text{'Manteiga'}\} \Rightarrow \{\text{'Pão'}\}$ (conf=1.00, sup=0.30)

$\{\text{'Café'}, \text{'Pão'}\} \Rightarrow \{\text{'Manteiga'}\}$ (conf=1.00, sup=0.30)

⇒ Totais: ItemSets1=3, ItemSets2=3, ItemSets3=1, Regras=7.