

2. Aprendizagem

a)

$$E(C|A) = \sum_{j=1}^M p(a_j) \times \left[- \sum_{i=1}^N p(c_i/a_j) \times \log_2(p(c_i/a_j)) \right]$$

C - conclusão

A - atributo

M - número de valores possíveis para o atributo A

N - número de valores possíveis para a conclusão C

$$E(\text{textura}) = \frac{8}{16} \left[- \frac{6}{8} \log_2\left(\frac{6}{8}\right) - \frac{2}{8} \log_2\left(\frac{2}{8}\right) \right]$$

pedado →

$$+ \frac{8}{16} \left[- \frac{4}{8} \log_2\left(\frac{4}{8}\right) - \frac{4}{8} \log_2\left(\frac{4}{8}\right) \right]$$

médio →

$$= 0,4056 + 0,5 = 0,9056 //$$

$$E(\text{cor}) = \frac{3}{16} \left[- \frac{3}{3} \log_2\left(\frac{3}{3}\right) - 0 \right]$$

castanho →

$$+ \frac{6}{16} \left[- \frac{2}{6} \log_2\left(\frac{2}{6}\right) - \frac{4}{6} \log_2\left(\frac{4}{6}\right) \right] + \frac{7}{16} \left[- \frac{5}{7} \log_2\left(\frac{5}{7}\right) - \frac{2}{7} \log_2\left(\frac{2}{7}\right) \right]$$

verde →

vermelho →

$$= 0 + 0,3444 + 0,3776 = 0,722 //$$

$$E(\text{tamanho}) = \frac{9}{16} \left[- \frac{5}{9} \log_2\left(\frac{5}{9}\right) - \frac{4}{9} \log_2\left(\frac{4}{9}\right) \right]$$

pequeno →

$$+ \frac{7}{16} \left[- \frac{5}{7} \log_2\left(\frac{5}{7}\right) - \frac{2}{7} \log_2\left(\frac{2}{7}\right) \right]$$

grande →

$$= 0,5575 + 0,3776 = 0,9351 //$$

$$E(\text{casca}) = \frac{6}{16} \left[- \frac{3}{6} \log_2\left(\frac{3}{6}\right) - \frac{3}{6} \log_2\left(\frac{3}{6}\right) \right]$$

mole →

$$+ \frac{10}{16} \left[- \frac{7}{10} \log_2\left(\frac{7}{10}\right) - \frac{3}{10} \log_2\left(\frac{3}{10}\right) \right]$$

dura →

$$= 0,375 + 0,5608 = 0,883 //$$

O atributo de menor entropia é cor → raiz da árvore

Caso 1: cor = castanho

$$E(\text{casca}) = \overset{\substack{\uparrow \text{dura} \\ \downarrow \text{mole}}}{\frac{2}{3} \times 0 + \frac{1}{3} \times 0} = 0$$

O atributo casca possui valor de entropia nulo \rightarrow é suficiente

Caso 1.1: cor = castanho e casca = dura \Rightarrow seguro

Caso 1.2: cor = castanho e casca = mole \Rightarrow seguro

Caso 2: cor = verde

$$E(\text{casca}) = \overset{\substack{\uparrow \text{dura} \\ \downarrow \text{mole}}}{\frac{3}{6} \left[-\frac{2}{3} \log_2 \left(\frac{2}{3} \right) - \frac{1}{3} \log_2 \left(\frac{1}{3} \right) \right]} + \frac{3}{6} \left[-\frac{1}{3} \log_2 \left(\frac{1}{3} \right) - \frac{2}{3} \log_2 \left(\frac{2}{3} \right) \right]$$
$$= 0,4591 + 0,4591 = 0,9182 //$$

$$E(\text{textura}) = \overset{\substack{\uparrow \text{peludo} \\ \downarrow \text{macio}}}{\frac{4}{6} \left[-\frac{2}{4} \log_2 \left(\frac{2}{4} \right) - \frac{2}{4} \log_2 \left(\frac{2}{4} \right) \right]}$$

$$+ \frac{2}{6} \times 0$$

$$= 0,6667 + 0 = 0,6667 //$$

$$E(\text{tamanho}) = \frac{4}{6} \times 0$$

$$+ \frac{2}{6} \times 0$$

$$= 0 //$$

O atributo tamanho possui valor de entropia nulo \rightarrow é suficiente

Caso 2.1: cor = verde e tamanho = pequeno \Rightarrow perigoso

Caso 2.2: cor = verde e tamanho = grande \Rightarrow seguro

Caso 3: cor = vermelho

$$E(\text{casca}) = \overset{\substack{\uparrow \text{dura} \\ \downarrow \text{mole}}}{\frac{5}{7} \left[-\frac{4}{5} \log_2 \left(\frac{4}{5} \right) - \frac{1}{5} \log_2 \left(\frac{1}{5} \right) \right]}$$

$$+ \frac{2}{7} \left[-\frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \log_2 \left(\frac{1}{2} \right) \right]$$

$$= 0,5157 + 0,2857 = 0,8014 //$$

$$E(\text{textura}) = \frac{3}{7} \times 0 + \frac{4}{7} \left[-\frac{2}{4} \log_2 \left(\frac{2}{4} \right) - \frac{2}{4} \log_2 \left(\frac{2}{4} \right) \right]$$

peludo ↗ macio ↗

$$= 0 + 0,5714 = 0,5714 //$$

$$E(\text{tamanho}) = \frac{4}{7} \times 0 + \frac{3}{7} \left[-\frac{1}{3} \log_2 \left(\frac{1}{3} \right) - \frac{2}{3} \log_2 \left(\frac{2}{3} \right) \right]$$

pequeno ↗ grande ↗

$$= 0 + 0,3936 = 0,3936 //$$

O valor de menor entropia é tamanho → nó da árvore

Caso 3.1: cor = vermelho e tamanho = pequeno ⇒ seguro

Caso 3.2:

$$E(\text{casca}) = \frac{1}{3} \times 0 + \frac{2}{3} \left[-\frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \log_2 \left(\frac{1}{2} \right) \right]$$

mole ↗ dura ↗

$$= 0 + 0,6667 = 0,6667 //$$

$$E(\text{textura}) = \frac{2}{3} \times 0 + \frac{1}{3} \times 0 = 0 //$$

macio ↗ peludo ↗

O atributo textura possui valor de entropia nulo → é suficiente

Caso 3.2.1: cor = vermelho e tamanho = grande e textura = macio ⇒ perigoso

Caso 3.2.2: cor = vermelho e tamanho = grande e textura = peludo ⇒ seguro

Árvore de decisão:

