



1. 计算 play 的信息熵。

$$\text{play} = \text{"no"} : 5.$$

$$\text{play} = \text{"yes"} : 9.$$

信息熵计算：

$$H(\text{play}) = - \sum_{i=1}^n p(x_i) \log_b(p(x_i))$$

$$= - \frac{5}{14} \log_2 \frac{5}{14} - \frac{9}{14} \log_2 \frac{9}{14} \approx 0.9403.$$

outlook 信息熵：

| | | | | |
|----------|----|----------------|-------------|----------------------------|
| overcast | 4 | $\frac{4}{14}$ | play = yes. | $\text{play} = \text{no.}$ |
| rainy | 5. | $\frac{5}{14}$ | 3. | 2 |
| sunny | 5. | $\frac{5}{14}$ | 2 | 3 |

$$H(\text{play} | \text{outlook}) = \frac{4}{14} (-\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4}) + \frac{5}{14} (-\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5}) + \frac{5}{14} (-\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5}) = 0.6935.$$

$$H(\text{play} | \text{temperature}) = 4/14 \times (-\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4}) + \frac{6}{14} \times (-\frac{4}{6} \log_2 \frac{4}{6} - \frac{2}{6} \log_2 \frac{2}{6}) + \frac{4}{14} \times (-\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4}) = 0.9111$$

$$H(\text{play} | \text{humidity}) = (\frac{7}{14} \times 0.9252 + \frac{7}{14} \times 0.5917) = 0.7885$$

$$H(\text{play} | \text{windy}) = \frac{8}{14} \times 0.8913 + \frac{6}{14} \times 1 = 0.8922$$

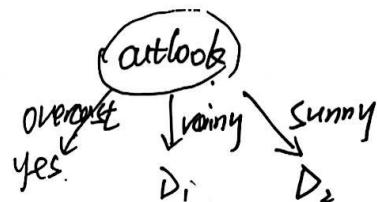
$$g(\text{play}, \text{outlook}) = 0.2467$$

构建模型结果：

$$g(\text{play}, \text{temperature}) = 0.0292$$

$$g(\text{play}, \text{humidity}) = 0.1518$$

$$g(\text{play}, \text{windy}) = 0.0481$$



$$D_1 \text{ 時} H(\text{play}) = -\frac{2}{5} \log_{\frac{1}{2}} - \frac{3}{5} \log_{\frac{1}{2}} = 0.9710$$

$$H(\text{play} | \text{temperature}) = \frac{2}{5} \times (-\frac{1}{2} \log_{\frac{1}{2}} - \frac{1}{2} \log_{\frac{1}{2}}) + \frac{3}{5} (-\frac{1}{3} \log_{\frac{1}{3}} - \frac{1}{3} \log_{\frac{1}{3}}) = 0.9710$$

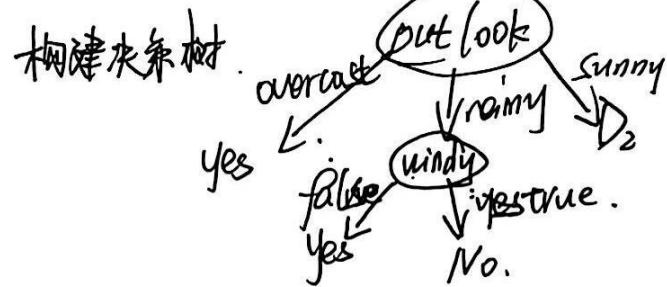
$$H(\text{play} | \text{humidity}) = \frac{3}{5} (-\frac{1}{3} \log_{\frac{1}{3}} - \frac{1}{3} \log_{\frac{1}{3}}) + \frac{2}{5} (-\frac{1}{2} \log_{\frac{1}{2}} - \frac{1}{2} \log_{\frac{1}{2}}) = 0.9510$$

$$H(\text{play} | \text{windy}) = \frac{2}{5} \times 0 + \frac{3}{5} \times 0 = 0$$

$$G(\text{play} | \text{temperature}) = 0.02$$

$$G(\text{play}, \text{humidity}) = 0.02$$

$$G(\text{play}, \text{windy}) = \underline{\underline{0.9710}}$$



$$D_2 \text{ 時} H(\text{play}) = -\frac{2}{5} \log_{\frac{1}{2}} - \frac{3}{5} \log_{\frac{1}{2}} = 0.9710$$

$$H(\text{play} | \text{temperature}) = \cancel{\frac{2}{5} \times 0} + \frac{2}{5} \times (-\frac{1}{2} \log_{\frac{1}{2}} - \frac{1}{2} \log_{\frac{1}{2}}) + \frac{1}{5} \times 0 = 0.4$$

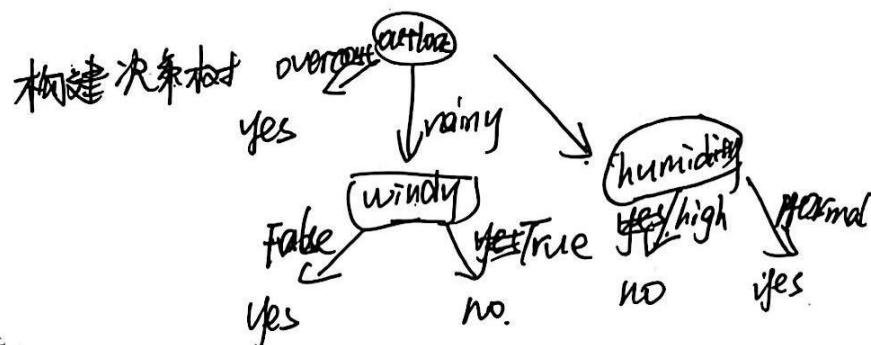
$$H(\text{play} | \text{humidity}) = \cancel{0.4} \frac{3}{5} \times 0 + \frac{2}{5} \times 0 = 0$$

$$H(\text{play} | \text{windy}) = \frac{2}{5} \times (-\frac{1}{2} \log_{\frac{1}{2}} - \frac{1}{2} \log_{\frac{1}{2}}) + \frac{3}{5} (-\frac{1}{3} \log_{\frac{1}{3}} - \frac{1}{3} \log_{\frac{1}{3}}) = 0.9510$$

$$g(\text{play} | \text{temperature}) = 0.5710$$

$$g(\text{play} | \text{humidity}) = 0.9710$$

$$g(\text{play} | \text{windy}) = 0.2$$



最終結果

