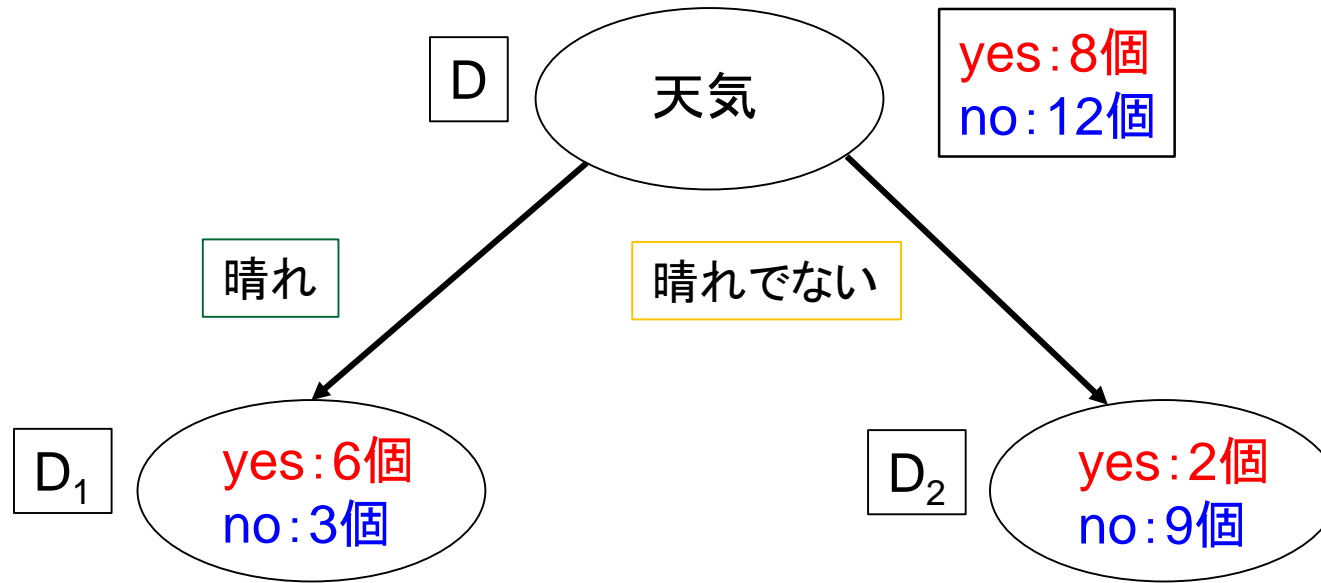


練習問題 回答(決定木)

管理工学科
篠沢 佳久

練習問題

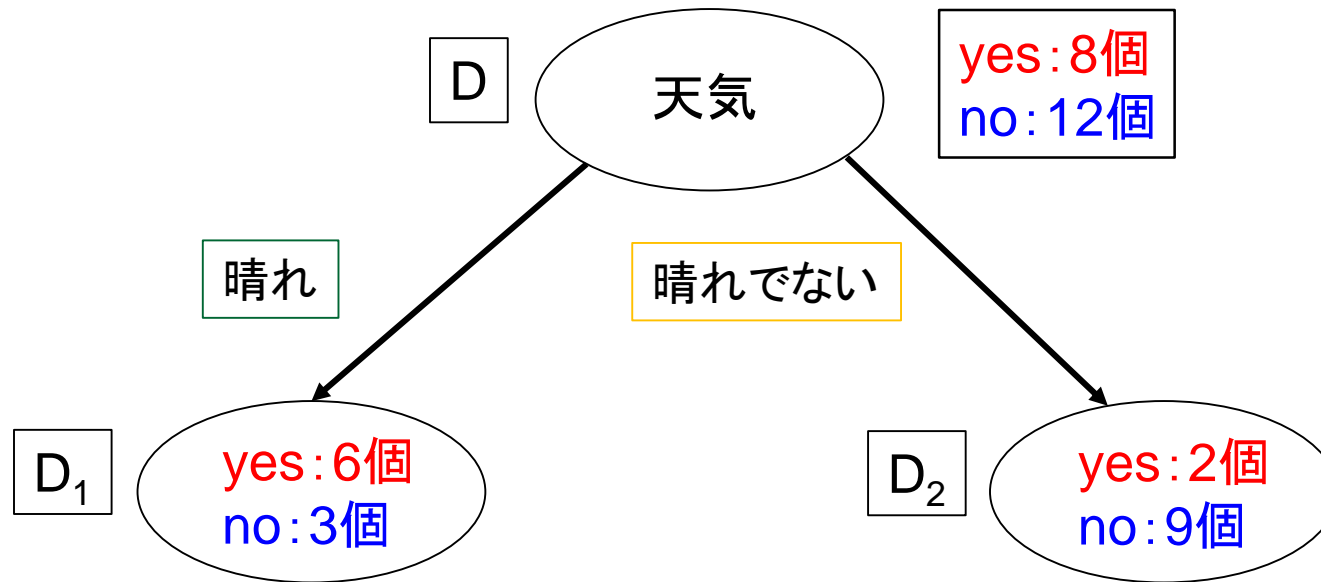


- ① 初期のエントロピーを求めなさい.
- ② D1のエントロピーを求めなさい.
- ③ D2のエントロピーを求めなさい.
- ④ 天気で分類した際のゲインを求めなさい.

対数の計算はExcelでして下さい*
=log(値, 2)

*特徴ごとにゲインの大小を比較するのが目的なので、対数の底は2でなくともかまいません

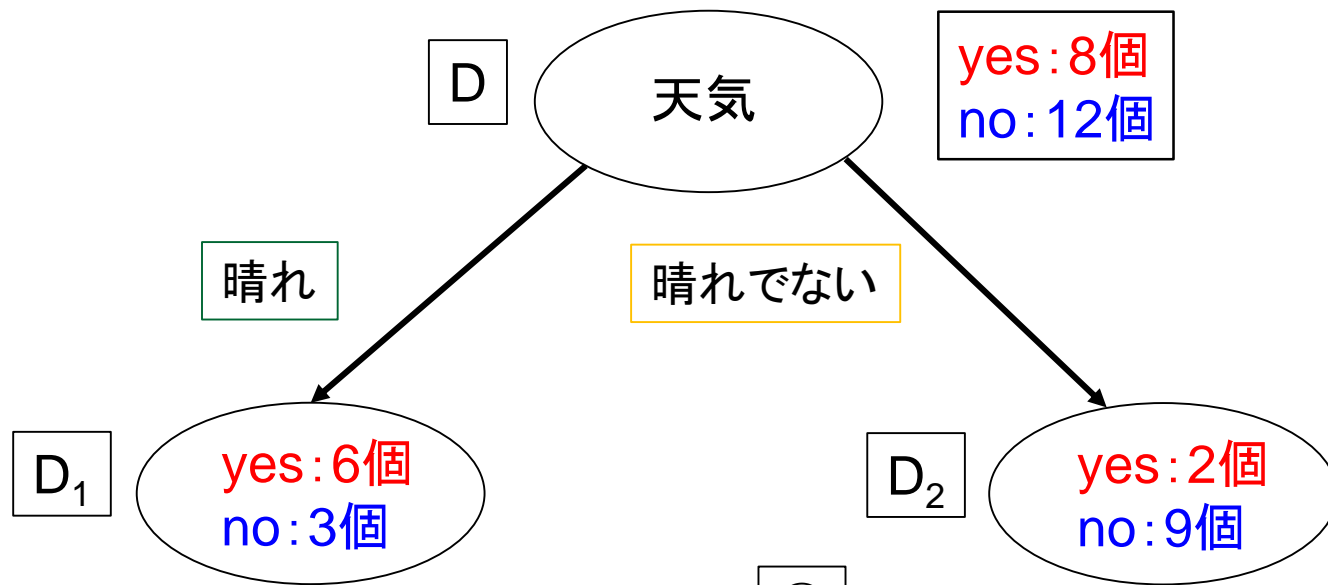
回答①



① 開始時のエントロピー

$$E(D) = -\frac{8}{20} \log \frac{8}{20} - \frac{12}{20} \log \frac{12}{20} = 0.971$$

回答②



②

$$E(D_1) = -\frac{6}{9} \log \frac{6}{9} - \frac{3}{9} \log \frac{3}{9} = 0.918$$

③

$$E(D_2) = -\frac{2}{11} \log \frac{2}{11} - \frac{9}{11} \log \frac{9}{11} = 0.684$$

④

$$\begin{aligned} \text{Gain}(D) &= E(D) - \frac{9}{20} E(D_1) - \frac{11}{20} E(D_2) \\ &= 0.971 - \frac{9}{20} \times 0.918 - \frac{11}{20} \times 0.684 = 0.181 \end{aligned}$$