

Exercise 3

$$\begin{aligned}\text{MI}(T_k, C_1) &= \log_2 \left(\frac{P[T_k, C_1]}{P[T_k]P[C_1]} \right) \\ &= \log_2 \left(\frac{\frac{10}{100}}{\frac{20}{100} \cdot \frac{45}{100}} \right) \\ &= \log_2 \left(\frac{10 \cdot 100}{45 \cdot 20} \right) \\ &\approx 0.152\end{aligned}$$

$$\begin{aligned}\text{MI}(T_k, C_2) &= \log_2 \left(\frac{P[T_k, C_2]}{P[T_k]P[C_2]} \right) \\ &= \log_2 \left(\frac{\frac{50}{165}}{\frac{55}{165} \cdot \frac{70}{165}} \right) \\ &= \log_2 \left(\frac{165 \cdot 50}{55 \cdot 70} \right) \\ &\approx 1.1\end{aligned}$$

$$\begin{aligned}\text{MI}(T_k, C_3) &= \log_2 \left(\frac{P[T_k, C_3]}{P[T_k]P[C_3]} \right) \\ &= \log_2 \left(\frac{\frac{100}{1100}}{\frac{300}{1100} \cdot \frac{400}{1100}} \right) \\ &= \log_2 \left(\frac{1100 \cdot 100}{300 \cdot 400} \right) \\ &\approx -0.126\end{aligned}$$

We can derive a category-independent term score by applying the sum function

$$f_{sum}(T_k) = \sum_{i=1}^{|C|} \text{MI}(T_k, C_i)$$

which will lead to the following result:

$$f_{sum}(T_k) = 0.152 + 1.1 - 0.126 = 1.126.$$