

习题二

15. (2) $x = (0, 1, 0, 1)$, $y = (1, 0, 1, 0)$. 计算余弦相似度、相关系数、欧氏距离、Jaccard 系数。

解: ① $\cos(x, y) = \frac{1 \times 0 + 1 \times 0 + 0 \times 1 + 0 \times 1}{\sqrt{0^2 + 1^2 + 0^2 + 1^2} \cdot \sqrt{1^2 + 0^2 + 1^2 + 0^2}} = 0$

② $\bar{x} = \frac{0+1+0+1}{4} = 0.5$, $\bar{y} = \frac{1+0+1+0}{4} = 0.5$

$\text{corr}(x, y) = \frac{-0.25 + (-0.25) + (-0.25) + (-0.25)}{\sqrt{0.5^2 + 0.5^2 + 0.5^2 + 0.5^2} \times \sqrt{0.5^2 + 0.5^2 + 0.5^2 + 0.5^2}} = -1$

③ $ob_2(x, y) = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2$

④ $J = \frac{n_{11}}{n_{01} + n_{10} + n_{11}} = \frac{0}{2+2} = 0$

18. (1) 计算简单匹配系数; (2) 计算杰卡德系数; (3) 哪两个人将成为最佳笔友? 哪两个最不相容?

解: (1) $k = (0, 1, 1, 0)$, $C = (0, 1, 1, 0)$, $E = (1, 0, 0, 1)$.

$\text{SMC}_{k,C} = \frac{n_{11} + n_{00}}{n_{11} + n_{00} + n_{01} + n_{10}} = \frac{2+2}{2+2+0+0} = \frac{4}{4} = 1$

同理, $\text{SMC}_{k,E} = \frac{0}{4} = 0$, $\text{SMC}_{C,E} = \frac{0}{4} = 0$

(2) $J_{k,C} = \frac{n_{11}}{n_{11} + n_{01} + n_{10}} = \frac{2}{2+0+0} = \frac{2}{2} = 1$

同理, $J_{k,E} = \frac{0}{4} = 0$, $J_{C,E} = \frac{0}{4} = 0$

(3) K和C两人的SMC与Jaccard系数最高, 均为1, 故可能成为最佳笔友; K与E、C与E的SMC与Jaccard系数均为最低, 为0, 故为最不相容的对。

(4) $J_{k,C} = \frac{2}{2}$, $J_{k,E} = \frac{1}{5}$, $J_{C,E} = \frac{0}{5} = 0$
仍为K和C两人。

22计科8班
陈煜祺
3122008883