Problema 1

Albert Ribes

24 de noviembre de 2017

Considerem un problema de classificació en dues classes, en les quals es disposa de les probabilitats de cada classe $P(C_1)$ i $P(C_2)$. Considerem tres possibles regles per classificar un objecte:

1. (R_1) Predir la classe més probable
2. (R_2) Predir la classe C_1 amb probabilitat $P(C_1)$
3. (R_3) Predir la classe C_1 amb probabilitat 0,5

Es demana:

1. Donar les probabilitats d'error $P_i(error)$ de les tres regles, $i = 1, 2, 3$
   - **El error de la regla R_1 es $min(P(C_1), P(C_2))$**
   - **Sea $Q(C_i)$ la probabilidad de predecir la clase C_i. La probabilidad de error de la regla R_2 es $P(C_1) \wedge Q(C_2) + P(C_2) \wedge Q(C_1)$. Puesto que P y Q son probabilidades independientes, se puede escribir como $P(C_1) \cdot Q(C_2) + P(C_2) \cdot Q(C_1)$. Pero $Q(C_i) = P(C_i)$, como indica la regla. Por lo tanto el error de la regla R_2 es $P(C_1) \cdot (1 - P(C_1)) + (1 - P(C_1)) \cdot P(C_1)$, que equivale a $2P(C_1) - 2P(C_1)^2$**
   - **Este es un caso particular de la regla R_2. Ahora el error es $P(C_1) \cdot 0,5 + (1 - P(C_1)) \cdot 0,5 \equiv 0,5$**

2. Demostrar que $P_1(error) \leq P_2(error) \leq P_3(error)$

   Sin pérdida de generalidad asumiremos que $P(C_1) \geq \frac{1}{2}$. El caso contrario es simétrico. Entonces para la primera parte de la demostración hay que demostrar que

$$\begin{aligned} P(C_2) \leq 2P(C_1) - 2P(C_1)^2 &\quad \equiv \\ 1 - P(C_1) \leq 2P(C_1) - 2P(C_1)^2 &\quad \equiv \\ 2P(C_1)^2 - 3P(C_1) + 1 \leq 0 & \end{aligned}$$

   Usaremos la fórmula de las ecuaciones de segundo grado para resolverlo:

$$P(C_1) = \frac{3 \pm \sqrt{9 - 4 \cdot 2 \cdot 1}}{2 \cdot 2} = \frac{3 \pm 1}{4} \Rightarrow P(C_1) \notin (\frac{1}{2}, 1)$$

   Esto es claramente una contradicción. Habrá que ver qué nos está pasando