Problema 1

Albert Ribes

10 de diciembre 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, $i = 1, 2, 3$

   Sea $Q(C_i)$ la probabilidad de elegir la clase C_i. En todos los casos, la probabilidad de error es $P_{error} = Q(C_1)P(C_2) + Q(C_2)P(C_1)$

   - **Para la regla R_1, si la clase más probable es C_1 la probabilidad de error será $P_{error} = 1 \times P(C_2) + 0 \times P(C_1) = P(C_2)$, y si la más probable es C_2 el error será $P_{error} = 0 \times P(C_2) + 1 \times P(C_1) = P(C_1)$. En cualquier caso, la probabilidad de error siempre será $P_{error} = min(P(C_1), P(C_2))$**
   - **Para la regla R_2 la probabilidad de error es**

     $$\begin{aligned} P_{error} &= P(C_1)P(C_2) & &+(1 - P(C_1))P(C_1) \\ &= P(C_1)(1 - P(C_1)) & &+(1 - P(C_1)P(C_1)) \\ &= 2P(C_1)(1 - P(C_1)) \\ &= 2P(C_1) - 2P(C_1)^2 \end{aligned}$$

   - **Para la regla R_3 la probabilidad de error es**

     $$\begin{aligned} P_{error} &= 0{,}5P(C_2) + 0{,}5P(C_1) \\ &= 0{,}5(1 - P(C_1)) + 0{,}5P(C_1) \\ &= 0{,}5 - 0{,}5P(C_1) + 0{,}5P(C_1) \\ &= 0{,}5 \end{aligned}$$

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

   Hay que demostrar que

$$
\begin{align}
min(P(C_1), P(C_2)) &\leq 2P(C_1) - 2P(C_1)^2 \tag{1}\\
2P(C_1) - 2P(C_1)^2 &\leq 0{,}5 \tag{2}
\end{align}
$$

En el caso que $P(C_1) \leq P(C_2)$, la condición 1 se puede escribir como

$$
\begin{aligned}
P(C_1) &\leq 2P(C_1) - 2P(C_1)^2\\
0 &\leq P(C_1) - 2P(C_1)^2\\
2P(C_1)^2 &\leq P(C_1)\\
2P(C_1) &\leq 1\\
P(C_1) &\leq \frac{1}{2}
\end{aligned}
$$

Y esto siempre es cierto, puesto que $P(C_1) + P(C_2) = 1$ y hemos establecido que $P(C_1) \leq P(C_2)$

En el caso que $P(C_1) > P(C_2)$, la condición 1 se puede escribir como

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

La igualdad se cumple en los puntos

$$
\begin{aligned}
x = \frac{3 \pm \sqrt{(-3)^2 - 4 \cdot 2}}{2 \cdot 2} &= \frac{3 \pm 1}{4}\\
x &\in \{\frac{1}{2}, 1\}
\end{aligned}
$$

Y como el elemento elevado al cuadrado es positivo, eso significa que:

$$P(C_1) \in [\frac{1}{2}, 1]$$

Igual que antes, esto es cierto puesto que hemos asumido que $P(C_1) > P(C_2)$.

La condición 1 ya está demostrada. Para la condición 2:

$$2P(C_1) - 2P(C_1)^2 \leq 0{,}5$$
$$0 \leq 2P(C_1)^2 - 2P(C_1) + 0{,}5$$

La igualdad se cumple en los puntos

$$x = \frac{2 \pm \sqrt{(-2)^2 - 4 \cdot 2 \cdot 0{,}5}}{2 \cdot 2} = \frac{2 \pm 0}{4}$$
$$x = \frac{1}{2}$$

Y como el elemento elevado al cuadrado es positivo, la inecuación siempre es cierta.

Quedan demostradas las dos condiciones