EJEMPLO DE CLASIFICACIÓN LINEAL USANDO SVM

Dados los siguientes vectores de soporte con un bias de entrada 1 (aumento de dimensión)

$\vec{S}_1$= (1, o, 1), $\vec{S}_2$= (3, 1, 1), $\vec{S}_3$= (3, -1, 1), los cuales están clasificados como sigue:

$\vec{S}_1$: clase negativa, $\vec{S}_2$: clase positiva, $\vec{S}_3$: clase positiva.

¿El punto (5,-1) es clasificado como negativo o positivo?

Nota.- Para clasificar un nuevo dato se usan los vectores soporte y el kernel de transformación (en este ejercicio el kernel es la función identidad) evaluando la función:

$$f(x) = \sigma \left(\textstyle\sum_i \alpha_i \, \Phi(\vec{S}_i).\Phi(x)\right)$$

SOLUCIÓN

En primer lugar, necesitamos hallar la ecuación del hiperplano que separe las clases: $\bar{w}\bar{x} + b=0$

Para obtener $\bar{w}$ y *b* debemos hacer uso de todo los vectores soporte y resolver el siguiente sistema de ecuaciones (una por cada vector soporte):

$\vec{S}_1$: clase negativa

$$\alpha_1 \, \Phi(\vec{S}_1).\Phi(\vec{S}_1) + \alpha_2 \, \Phi(\vec{S}_2).\Phi(\vec{S}_1) + \alpha_3 \Phi(\vec{S}_3).\Phi(\vec{S}_1) = -1 \quad \ldots\ldots. (1)$$

$\vec{S}_2$: clase positiva

$$\alpha_1 \, \Phi(\vec{S}_1).\Phi(\vec{S}_2) + \alpha_1 \, \Phi(\vec{S}_2).\Phi(\vec{S}_2) + \alpha_1 \, \Phi(\vec{S}_3).\Phi(\vec{S}_2) = +1 \quad \ldots\ldots(2)$$

$\vec{S}_3$: clase positiva

$$\alpha_1 \, \Phi(\vec{S}_1).\Phi(\vec{S}_3) + \alpha_1 \, \Phi(\vec{S}_2).\Phi(\vec{S}_3) + \alpha_1 \, \Phi(\vec{S}_3).\Phi(\vec{S}_3) = +1 \quad \ldots\ldots(3)$$

El siguiente conjunto de ecuaciones, se obtiene al optimizar con respecto a los multiplicadores de Lagrange, para lo cual debemos resolverlos usando los coeficientes α. Como en estas ecuaciones estamos usando el kernel, debemos aplicar la función de transformación definida con él.

Al usar un **kernel lineal** (Φ) que es la función identidad con un bias adicional (x_0=1), el sistema de ecuaciones se reduce a:

$$\alpha_1 \, \vec{S}_1.\vec{S}_1 + \alpha 2 \, \vec{S}_2.\vec{S}_1 + \alpha 3 \, \vec{S}_3.\vec{S}_1 = -1 \quad \ldots\ldots. (1)$$

$$\alpha_1 \, \vec{S}_1.\vec{S}_2 + \alpha 2 \, \vec{S}_2.\vec{S}_2 + \alpha 3 \, \vec{S}_3.\vec{S}_2 = +1 \quad \ldots\ldots. (2)$$

$$\alpha_1 \, \vec{S}_1.\vec{S}_3 + \alpha 2 \, \vec{S}_2.\vec{S}_3 + \alpha 3 \, \vec{S}_3.\vec{S}_3 = +1 \quad \ldots\ldots. (3)$$

Efectuando el producto escalar, tenemos:

$$\begin{aligned} 2\alpha_1 + 4\alpha_2 + 4\alpha_3 &= -1 \\ 4\alpha_1 + 11\alpha_2 + 9\alpha_3 &= +1 \\ 4\alpha_1 + 9\alpha_2 + 11\alpha_3 &= +1 \end{aligned}$$

Desarrollando el sistema de ecuaciones, obtenemos α_1 = -3.5, α_2 = 0.75 y α_3 = 0.75,

Para clasificar un nuevo dato se usan los vectores soporte y el kernel de transformación:

$$f(x) = \sigma \left(\textstyle\sum_i \alpha_i \, \Phi(\vec{S}_i).\Phi(x)\right)$$

Por lo tanto:

$$f(5,-1) = \sigma\left[-3.5\,\Phi\begin{bmatrix}1\\0\end{bmatrix}.\,\Phi\begin{bmatrix}5\\-1\end{bmatrix} + 0.75\,\Phi\begin{bmatrix}3\\1\end{bmatrix}.\,\Phi\begin{bmatrix}5\\-1\end{bmatrix} + 0.75\,\Phi\begin{bmatrix}3\\-1\end{bmatrix}.\,\Phi\begin{bmatrix}5\\-1\end{bmatrix}\right]$$

$$= \sigma\left[-3.5\begin{bmatrix}1\\0\\1\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix} + 0.75\begin{bmatrix}3\\1\\1\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix} + 0.75\begin{bmatrix}3\\-1\\1\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix}\right]$$

$$= \sigma\left[\begin{bmatrix}-3.5\\0\\-3.5\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix} + \begin{bmatrix}2.25\\0.75\\0.75\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix} + \begin{bmatrix}2.25\\-0.75\\0.75\end{bmatrix}.\begin{bmatrix}5\\-1\\1\end{bmatrix}\right]$$

= σ (-23.0 + 11.25 + 12.75)
= σ (**1.0**)

Por lo tanto (5,-1) es clasificado como **POSITIVO**