UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Algoritmo Perceptrón[1]

Jorge Civera
Alfons Juan
Albert Sanchis

DSIC
Departamento de Sistemas
Informáticos y Computación

[1] Para una correcta visualización, se requiere Acrobat Reader v. 7.0 o superior

Objetivos formativos

- Aplicar el algoritmo Perceptrón a una tarea de clasificación
- Explicar el comportamiento del algoritmo Perceptrón en función de sus parámetros

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Índice

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

1. Funciones discriminantes lineales

Todo clasificador puede representarse como:

$$c(x) = \arg\max_{c} \; g_c(x)$$

donde cada clase c utiliza una ***función discriminante*** $g_c(x)$ que mide el grado de pertenencia de un objeto x a la clase c

Las funciones discriminantes más utilizadas son ***lineales*** (con $\boldsymbol{x}$):

$$g_c(\boldsymbol{x}) = \boldsymbol{w}_c^t \boldsymbol{x} + w_{c0} \qquad \text{donde} \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_D \end{pmatrix} \quad \text{y} \quad \boldsymbol{w_c} = \begin{pmatrix} w_{c1} \\ \vdots \\ w_{cD} \end{pmatrix}$$

Con notación ***homogénea***:

$$g_c(\mathbf{x}) = \mathbf{w}_c^t \mathbf{x} \qquad \text{donde} \quad \mathbf{x} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \quad \text{y} \quad \mathbf{w}_c = \begin{pmatrix} w_{c0} \\ \boldsymbol{w}_c \end{pmatrix}$$

2. Algoritmo Perceptrón

Entrada: $\{(\mathbf{x}_n, c_n)\}_{n=1}^{N}, \quad \{\mathbf{w}_c\}_{c=0}^{C}, \quad \alpha \in \mathbb{R}^{>0}$ **y** $b \in \mathbb{R}$

Salida: $$\{\mathbf{w}_c\}^* = \arg\min_{\{\mathbf{w}_c\}} \sum_n \left[\max_{c \neq c_n} \mathbf{w}_c^t \mathbf{x}_n + b > \mathbf{w}_{c_n}^t \mathbf{x}_n \right]$$

Método:

$$[\,P\,] = \begin{cases} 1 & \text{si} \quad P = \text{verdadero} \\ 0 & \text{si} \quad P = \text{falso} \end{cases}$$

repetir

para todo dato $\mathbf{x}_n$

$err = \texttt{falso}$

para toda clase c distinta de c_n

si $\mathbf{w}_c^t \mathbf{x}_n + b > \mathbf{w}_{c_n}^t \mathbf{x}_n$: $\mathbf{w}_c = \mathbf{w}_c - \alpha \cdot \mathbf{x}_n$; $err = \texttt{verdadero}$

si err: $\mathbf{w}_{c_n} = \mathbf{w}_{c_n} + \alpha \cdot \mathbf{x}_n$

hasta que no quedan muestras mal clasificadas

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

3. Ejemplo

4. Convergencia y calidad de la solución

Converge si los datos son linealmente separables y $b \leq 0$

Conviene implementarlo con un máximo número de iteraciones.

Cuando $\alpha \to 0$, la convergencia es más suave, pero más lenta.

Calidad de la solución:

Linealmente separables	$b \leq 0$	$b > 0$
SI	Fronteras con poca holgura	Fronteras *centradas*
NO	Fronteras baja calidad	Fronteras casi óptimas

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

5. Conclusiones

Hemos visto:

- El algoritmo Perceptrón y una traza del mismo
- La convergencia del algoritmo en función de sus parámetros y las muestras de entrenamiento utilizadas

UNIVERSITAT POLITÈCNICA DE VALÈNCIA