

CLASIFICADORES BASADOS EN INSTANCIAS

Ejercicio 1

X	0.5	3.0	4.5	4.6	4.9	5.2	5.3	5.5	7.0	9.5
Y	-	-	+	+	+	-	-	+	-	-

Apartado a. Clasifica el punto $x=5$ de acuerdo a los 1-,3-,5- y 9- vecinos más cercanos usando mayoría de votos

Calculamos todas las instancias desde $x=5$

Distancias	Sorted	Y
$D(5,0.5) = 4.5$	$D(5,4.9) = 0.1$	+
$D(5,0.5) = 4.5$	$D(5,5.2) = 0.2$	-
$D(5,4.5) = 0.5$	$D(5,5.3) = 0.3$	-
$D(5,4.6) = 0.4$	$D(5,4.6) = 0.4$	+
$D(5,4.9) = 0.1$	$D(5,4.5) = 0.5$	+
$D(5,5.2) = 0.2$	$D(5,5.5) = 0.5$	+
$D(5,5.3) = 0.3$	$D(5,3) = 2$	-
$D(5,5.5) = 0.5$	$D(5,7) = 2$	-
$D(5,7) = 2$	$D(5,0.5) = 4.5$	-
$D(5,9.5) = 4.5$	$D(5,9.5) = 4.5$	-

Obtenemos la clase para los 1-, 3- 5- y 9- vecinos más cercanos

1- vecino más cercano $\rightarrow d(5,4.9) \rightarrow$ Clase positiva

3- vecino más cercano $\rightarrow [d(5,4.9),d(5,5.3)] \rightarrow$ Clase negativa

5- vecino más cercano $\rightarrow [d(5,4.9),d(5,4.5)] \rightarrow$ Clase positiva

9- vecino más cercano $\rightarrow [d(5,0.5), d(5,9.5)] \rightarrow$ Clase negativa

Apartado b. Repetir el apartado anterior usando:

$$y = \underset{(x_i, y_i) \in D_z}{\operatorname{argmax}} \sum (W_i * I(V = y_i)), W_i = \frac{1}{d(X', X_i)^2}$$

Calculamos las distancias desde X=5

Distancia	Sorted	Clase
D(5,0.5) = 0.049	D(5,4.9) = 100	+
D(5,3.0) = 0.25	D(5,5.2) = 2.5	-
D(5,4.5) = 4	D(5,5.3) = 11.111	-
D(5,4.6) = 6.250	D(5,4.6) = 6.25	+
D(5,4.9) = 100	D(5,4.5) = 4	+
D(5,5.2) = 25	D(5,5.5) = 4	+
D(5,5.3) = 11.111	D(5,3) = 0.25	-
D(5,7) = 0.16	D(5,0.5) = 0.049	-
D(5,9.5) = 0.049	D(5,9.5) = 0.049	-

Obtenemos la clase para los vecinos indicados en el ejercicio:

1-nearest neighbor → d(5,4.9) = 100 → Clase positiva

3-nearest neighbor → [d(5,4.9), d(5,5.3)] = 100 → Clase positiva

5-nearest neighbor → [d(5,4.9), d(5,4.5)] = 100 → Clase positiva

9-nearest neighbor → [d(5,4.9), d(5,9.5)] = 100 → Clase positiva

Ejercicio 2. Calcula el MVDM measure para (Home Owner, Marital Status)

$$d(V_1, V_2) = \sum_{i=1}^K \left| \frac{ni_1}{n_1} - \frac{ni_2}{n_2} \right|$$

TID	HOME OWNER	MARITAL STATUS	ANNUAL INCOME	DEFAULT BORROWER
1	YES	SINGLE	125 K	NO
2	NO	MARITED	100 K	NO
3	NO	SINGLE	70 K	NO
4	YES	MARRIED	120 K	NO
5	NO	DIVORCED	95 K	YES
6	NO	MARRIED	60 K	NO
7	YES	DIVORCED	220 K	NO
8	NO	SINGLE	85 K	YES
9	NO	MARRIED	75 K	NO
10	NO	SINGLE	90 K	YES

Obtenemos el número de cada atributo de Marital Status y Home Owner en función del valor del atributo Default Borrower:

	MARITAL STATUS			HOME OWNER	
	Single	Married	Divorced	Yes	No
YES	2	0	1	0	3
No	2	4	1	3	4

Calculamos las distancias:

$d(\text{Single}, \text{Married}) = 1$; $d(\text{single}, \text{divorced}) = 0$; $d(\text{married}, \text{divorced}) = 1$

$d(\text{Refund} = \text{yes}, \text{Refund} = \text{no}) = \frac{6}{7}$