Tema 5. Clasificadores K-NN

Abdelmalik Moujahid, Iaki Inza y Pedro Larraaga
Departamento de Ciencias de la Computacin e Inteligencia Artificial
Universidad del Pas Vasco–Euskal Herriko Unibertsitatea

1. La tabla adjunta contiene 8 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN. Los dos últimos casos -numerados del 9 al 10- forman el conjunto de testeo.

<i>Caso</i>	<i>X</i>	<i>Y</i>	<i>Clase</i>
1	1	1	0
2	2	4	0
3	3	2	0
4	3	5	0
5	4	4	0
6	4	7	1
7	6	4	1
8	6	6	1
9	2	7	?
10	3	4	?

Construye el clasificador k -NN a partir de la distancia euclídea:

- Con el algoritmo k -NN básico ($k = 1, 2$).
 - Con el k -NN con rechazo con $k= 4$.
 - Con el k -NN con distancia media con $k= 4$.
 - Con el k -NN con distancia mínima.
 - Con el k -NN con pesado de casos seleccionados, siendo el mismo inversamente proporcional a la distancia. $k= 4$.
 - Estudia los problemas que se plantean en este ejercicio si se quiere pesar las variables a partir de la cantidad de información mútua entre cada variable predictora y la variable clase.
 - Con una reducción del fichero inicial en base al método denominado edición de Wilson. $k=1$.
 - Con una reducción del fichero inicial en base al método denominado condensación de Hart. $k=1$.
2. La tabla adjunta contiene 8 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN.

<i>Caso</i>	<i>X₁</i>	<i>X₂</i>	<i>Clase</i>
1	2	0	0
2	4	4	1
3	1	1	0
4	2	4	1
5	2	2	0
6	2	3	1
7	3	4	0
8	3	3	1

- Aplica la edición de Wilson para reducir el conjunto de entrenamiento.
 - Usando el método de condensación de Hart, reduce el conjunto de entrenamiento.
 - Clasifica el caso (2.5, 2.5) a partir de los conjuntos de entrenamiento reducidos obtenidos anteriormente y considerando $k = 1$.
3. La tabla adjunta contiene 6 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN. Los dos últimos casos -numerados como 7 y 8- forman el conjunto de testeo.

<i>Caso</i>	<i>X</i>	<i>Y</i>	<i>Clase</i>
1	1	1	0
2	2	4	0
3	3	2	0
4	3	5	0
5	4	4	0
6	4	7	1
7	6	4	?
8	6	6	?

Construye el clasificador k -NN a partir de la distancia euclídea y de dos valores distintos de k ($k = 1, 2$).

4. La tabla adjunta contiene 8 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN. Los cuatro últimos casos -numerados del 9 al 12- constituyen el conjunto de testeo.

<i>Caso</i>	<i>X</i>	<i>Y</i>	<i>Clase</i>
1	1	1	0
2	2	4	0
3	3	2	0
4	3	5	0
5	4	4	0
6	4	7	1
7	6	4	1
8	6	6	1
9	2	7	?
10	3	4	?
11	4	6	?
12	5	5	?

Construye el clasificador k -NN a partir de la distancia euclídea y de dos valores distintos de k ($k = 1, 2$).