

2. K-NN sailkatzaileak Ariketak

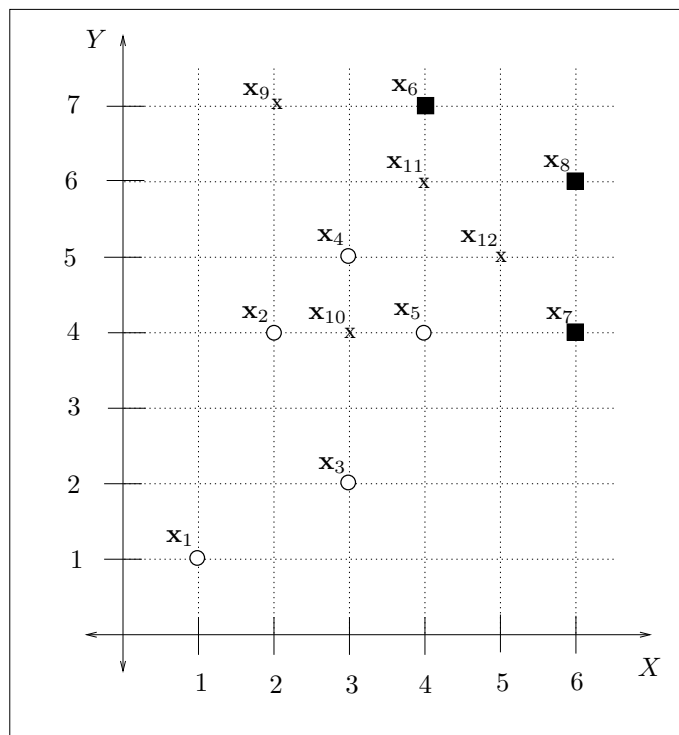
- Ondorengo taulan k -NN sailkatzaile bat eraikitzeko entrenamendurako kasuen multzoa agertzen da. Bertan bi dimentsiotako 8 kasu daude. Taulako azken lau kasuek testerako multzoa osatzen dute.

Kasua	X	Y	C		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1	1	0	x_1	0	3.16	2.24	4.47	4.24	6.71	5.83	7.07
x_2	2	4	0	x_2	3.16	0	2.24	1.41	2.00	3.61	4.00	4.47
x_3	3	2	0	x_3	2.24	2.24	0	3.00	2.24	5.10	3.61	5.00
x_4	3	5	0	x_4	4.47	1.41	3.00	0	1.41	2.24	3.16	3.16
x_5	4	4	0	x_5	4.24	2.00	2.24	1.41	0	3.00	2.00	2.83
x_6	4	7	1	x_6	6.71	3.61	5.10	2.24	3.00	0	3.61	2.24
x_7	6	4	1	x_7	5.83	4.00	3.61	3.16	2.00	3.61	0	2.00
x_8	6	6	1	x_8	7.07	4.47	5.00	3.16	2.83	2.24	2.00	0
x_9	2	7	?	x_9	6.08	3.00	5.10	2.24	3.61	2.00	5.00	4.12
x_{10}	3	4	?	x_{10}	3.61	1.00	2.00	1.00	1.00	3.16	3.00	3.61
x_{11}	4	6	?	x_{11}	5.83	2.83	4.12	1.41	2.00	1.00	2.83	2.00
x_{12}	5	5	?	x_{12}	5.66	3.16	3.61	2.00	1.41	2.24	1.41	1.41

Datubasea

Datubaseko kasuen arteko distantzia Euklidentarrak

Datubasearen adierazpen grafikoa espazio Euklidentarrean:



Distantzia Euklidestarra erabiliz k -NN sailkatzailea eraiki ezazu:

- 1.1 Oinarrizko k -NN algoritmoa erabiliz, $k = 1$ eta $k = 2$ balioetarako.
- 1.2 k -NN bermegabeak baztertuz aldaera erabiliz, $k = 4$ eta atalasea 3 izanik.
- 1.3 k -NN batazbesteko distantziarekin aldaera erabiliz eta $k = 4$.
- 1.4 k -NN distantzia minimoarekin aldaera erabiliz, $k = 1$ kasurako (klase bakoitzerako barizentritik hurbilen dagoen kasua aukeratu ordezkari gisa).
- 1.5 k -NN aukeratutako auzokideen pisaketarekin aldaera erabiliz. Pisua distantziaren alderantzizko proportzionala erabiliz kalkula ezazu eta $k = 4$.
- 1.6 k -NN aldagaien pisaketarekin aldaera erabiliz. Aldagaien pisua kalkulatzeko aldagai iragarle bakoitzaren eta klase-aldagaiaren arteko elkarrekiko informazioaren neurria erabil ezazu, $k = 1$.
- 1.7 Wilson-en edizioa erabiliz murriztu ezazu 8 kasuko entrenamendurako kasuen multzoa, $k = 1$.
- 1.8 Hart-en kondentsazioa erabiliz murriztu ezazu 8 kasuko entrenamendurako kasuen multzoa, $k = 1$.

2. Aurreko ataleko galdera berberak erantzun itzazu entrenamendurako ondorengo 8 kasuak eta testerako kasu bakarra dituen datubase honetarako.

Kasua	X_1	X_2	C		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	2	0	0	x_1	0	4.47	1.41	4.00	2.00	3.00	4.12	3.16
x_2	4	4	1	x_2	4.47	0	4.24	2.00	2.83	2.24	1.00	1.41
x_3	1	1	0	x_3	1.41	4.24	0	3.16	1.41	2.24	3.61	2.83
x_4	2	4	1	x_4	4.00	2.00	3.16	0	2.00	1.00	1.00	1.41
x_5	2	2	0	x_5	2.00	2.83	1.41	2.00	0	1.00	2.24	1.41
x_6	2	3	1	x_6	3.00	2.24	2.24	1.00	1.00	0	1.41	1.00
x_7	3	4	0	x_7	4.12	1.00	3.61	1.00	2.24	1.41	0	1.00
x_8	3	3	1	x_8	3.16	1.41	2.83	1.41	1.41	1.00	1.00	0
x_9	2.5	2.5	?	x_9	2.55	2.12	2.12	1.58	0.71	0.71	1.58	0.71

Datubasea

Datubaseko kasuen arteko distantzia Euklidentarrak

Datubasearen adierazpen grafikoa espazio Euklidentarrean:

