

Zaawansowane uczenie maszynowe: *ćwiczenia do wykładu 14*

Paweł Cichosz

x	a_1	a_2	a_3	c	f
1	1	1	2	0	2
2	1	2	3	0	5
3	1	3	4	0	8
4	2	1	1	0	3
5	2	2	2	0	4
6	1	1	1	1	1
7	1	2	1	1	3
8	1	3	4	1	6
9	2	2	3	1	7
10	2	3	2	1	5

1. Traktując atrybuty jako numeryczne, wyznaczyć podzbiór dwóch atrybutów o najwyższych wartościach bezwzględnych współczynnika korelacji liniowej z f .
2. Traktując atrybuty jako numeryczne, wyznaczyć podzbiór dwóch atrybutów o najwyższym stosunku średniej wartości bezwzględnej ich współczynnika korelacji liniowej z f do wartości bezwzględnej współczynnika korelacji liniowej między nimi.
3. Traktując atrybuty jako dyskretne, wyznaczyć podzbiór dwóch atrybutów o najwyższych wartościach informacji wzajemnej z c .
4. Traktując atrybuty jako dyskretne, wyznaczyć podzbiór dwóch atrybutów o najwyższym stosunku średniej wartości ich informacji wzajemnej z c do wartości informacji wzajemnej między nimi.
5. Wyznaczyć parametry standardyzacji oraz normalizacji atrybutów na podstawie pierwszych ośmiu przykładów i zastosować je do transformacji wartości atrybutów dla pozostałych dwóch przykładów.
6. Wyznaczyć macierz transformacji PCA redukującej wymiarowość do 2 składowych głównych, posługując się wybranymi narzędziami obliczeń numerycznych w celu realizacji operacji wyznaczania macierzy kowariancji, wyznaczania wartości i wektorów własnych oraz mnożenia macierzy.

7. Wyznaczyć podział zakresu wartości atrybutu a_3 na przedziały minimalizujący entropię c .