

تمرین ۳

Euclidean distance: $d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2}$.

$i = (x_{i1}, x_{i2}, \dots, x_{in})$, and $j = (x_{j1}, x_{j2}, \dots, x_{jn})$,

$$\sqrt{(22 - 20)^2 + (1 - 0)^2 + (42 - 36)^2 + (10 - 8)^2} = \sqrt{45} = 6.7082$$

Manhattan distance: $d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{in} - x_{jn}|$.

$$|22 - 20| + |1 - 0| + |42 - 36| + |10 - 8| = 11.$$

Minkowski distance: $d(i, j) = \sqrt[h]{|x_{i1} - x_{j1}|^h + |x_{i2} - x_{j2}|^h + \dots + |x_{ip} - x_{jp}|^h}$

$$\sqrt[3]{|22 - 20|^3 + |1 - 0|^3 + |42 - 36|^3 + |10 - 8|^3} = \sqrt[3]{233} = 6.1534.$$

supremum distance: $d(i, j) = \lim_{h \rightarrow \infty} (\sum_{f=1}^p |x_{if} - x_{jf}|^h)^{1/h} = \max_f |x_{if} - x_{jf}|$

$$d(i, j) = 6$$

	Euclidean dist.	Manhattan dist	supremum dist	cosine sim
X1	0.1414	0.2	0.1	0.99999
X2	0.6708	0.9	0.6	0.99575
X3	0.2828	0.4	0.2	0.99997
X4	0.2236	0.3	0.2	0.99903
X5	0.6083	0.7	0.6	0.96536

بر اساس مقادیر به دست آمده در جدول بالا (بر اساس فرمول های ذکر شده در تمرین قبل) رنکینگ نقاط بر پایه شباهت به این صورت خواهد بود:

Euclidean distance: x1, x4, x3, x5, x2

Manhattan distance: x1, x4, x3, x5, x2

Supremum distance: x1, x4, x3, x5, x2

Cosine similarity: x1, x3, x4, x2, x5

کوثری نرمالزه شده برابر است با: (0.65850, 0.75258)

و مجموعه داده های نرمالیزه شده به این صورت خواهد بود:

	A1	A2
X1	0.66162	0.74984
X2	0.72500	0.68875
X3	0.66436	0.74741
X4	0.62470	0.78087
X5	0.83205	0.55470

با محاسبه مجدد Euclidean distances داریم:

	Euclidean distance
X1	0.00415
X2	0.09217
X3	0.00781
X4	0.04409
X5	0.26320

رنکینگ نهایی به این صورت به دست می آید:

x1, x3, x4, x2, x5