

Distance Measures

Question 1:

Consider the following three vectors u, v, w in a 6-dimensional space:

$$u = [1, 0.25, 0, 0, 0.5, 0]$$

$$v = [0.75, 0, 0, 0.2, 0.4, 0]$$

$$w = [0, 0.1, 0.75, 0, 0, 1]$$

Suppose $\cos(x,y)$ denotes the similarity of vectors x and y under the cosine similarity measure. Compute all three pairwise similarities among u, v, w .

1) Answer:
Given data is:

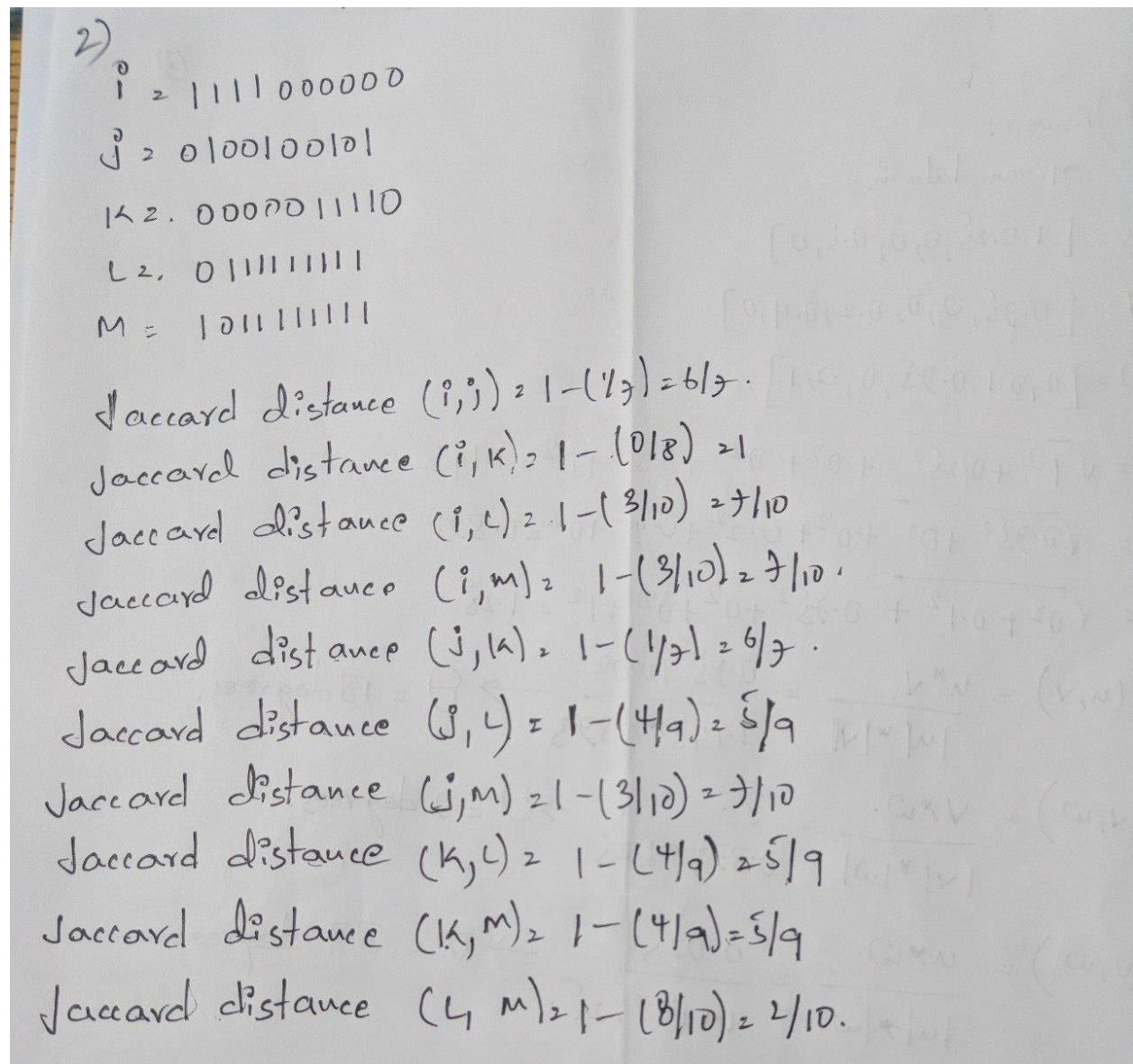
$$u = [1, 0.25, 0, 0, 0.5, 0]$$
$$v = [0.75, 0, 0, 0.2, 0.4, 0]$$
$$w = [0, 0.1, 0.75, 0, 0, 1]$$
$$|u| = \sqrt{1^2 + 0.25^2 + 0^2 + 0^2 + 0.5^2 + 0^2} = 1.145$$
$$|v| = \sqrt{0.75^2 + 0^2 + 0^2 + 0.2^2 + 0.4^2 + 0^2} = 0.873$$
$$|w| = \sqrt{0^2 + 0.1^2 + 0.75^2 + 0^2 + 0^2 + 1^2} = 1.25$$
$$\cos(u, v) = \frac{u \cdot v}{|u| \cdot |v|} = \frac{0.75 + 0.02}{1.145 \cdot 0.873} \rightarrow \theta = 18 \text{ degrees.}$$
$$\cos(v, w) = \frac{v \cdot w}{|v| \cdot |w|} = \frac{0}{0.873 \cdot 1.25} \rightarrow \theta = 90 \text{ degrees.}$$
$$\cos(u, w) = \frac{u \cdot w}{|u| \cdot |w|} = \frac{0.025}{1.145 \cdot 1.25} \rightarrow \theta = 89 \text{ degrees.}$$

Question 2:

Here are five vectors in a 10-dimensional space:

1111000000 0100100101 0000011110 0111111111 1011111111

Compute the Jaccard distance (not Jaccard "measure") between each pair of the vectors.



Question 3:

Here are five vectors in a 10-dimensional space:

1111000000 0100100101 0000011110 0111111111 1011111111

Compute the Manhattan distance (L_1 norm) between each two of these vectors.

3) a.

Let $i = 1111000000$, $j = 0100100101$, $k = 0000011110$

$L = 01111111$, $M = 10111111$

Manhattan distance of $i, j = 6$.

"	"	"	$i, k = 8$.
"	"	"	$i, L = 7$.
"	"	"	$i, M = 8$.
"	"	"	$j, k = 6$.
"	"	"	$j, L = 5$.
"	"	"	$j, M = 7$.
"	"	"	$k, L = 5$.
"	"	"	$k, M = 5$.
"	"	"	DEF $L, M = 2$.

Question 4: The edit distance is the minimum number of character insertions and character deletions required to turn one string into another. Compute the edit distance between each pair of the strings **he**, **she**, **his**, and **hers**.

4) The edit distance between. he and she = 1
 " " " " he and his = 2
 " " " " he and hers = 2.
 " " " " she and his = 4
 " " " " she and hers = 3.
 " " " " his and hers = 3.