### **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.

**Given Vectors:**

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]

cos (x, y) = =

cos (u, v) =

Therefore the equation is:

=

= **0.9503**

cos (u, w) =

Therefore the equation is:

=

**0.1742**

cos (v, w) =

Therefore the equation is:

=

= **0**

**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.

**Given Vectors:**

V1 = 1111000000

V2 = 0100100101

V3 = 0000011110

V4 = 0111111111

V5 = 1011111111

Jaccard similarity (A, B) =

Jaccard distance = 1 – Jaccard Similarity

* Jaccard similarity (V1, V2) = =

Jaccard distance (V1, V2) = 1 – =

* Jaccard similarity (V1, V3) = =

Jaccard distance (V1, V3) = 1 – 0=

* Jaccard similarity (V1, V4) = =

Jaccard distance (V1, V4) = 1 – =

* Jaccard similarity (V1, V5) = =

Jaccard distance (V1, V5) = 1 – =

* Jaccard similarity (V2, V3) = =

Jaccard distance (V2, V2) = 1 – =

**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.

**Given Vectors:**

V1 = 1111000000

V2 = 0100100101

V3 = 0000011110

V4 = 0111111111

V5 = 1011111111

We know that Manhattan distance is the sum of absolute difference of the components of the vectors.

Manhattan distance (x, y) = | – | + | – | + . . . + | – |

* Manhattan distance (V1, V2)

= |1-0|+|1-1|+|1-0|+|1-0|+|0-1|+|0-0|+|0-0|+|0-1|+|0-0|+|0-1| = 6

* Manhattan distance (V1, V3)

= 1+1+1+1+0+1+1+1+1+0 = 8

* Manhattan distance (V1, V4)

= 1+0+0+0+1+1+1+1+1+1 = 7

* Manhattan distance (V1, V5)

= 0+1+0+0+1+1+1+1+1+1 = 7

* Manhattan distance (V2, V3)

= 0+1+0+0+1+1+1+0+1+1 = 6

* Manhattan distance (V2, V4)

= 0+0+1+1+0+1+1+0+1+0 = 5

* Manhattan distance (V2, V5)

= 1+1+1+1+0+1+1+0+1+0 = 7

* Manhattan distance (V3, V4)

= 0+1+1+1+1+0+0+0+0+1 = 5

* Manhattan distance (V3, V5)

= 1+0+1+1+1+0+0+0+0+1 = 5

* Manhattan distance (V4, V5)

= 1+1+0+0+0+0+0+0+0+0 = 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**.

Given words are **he**, **she**, **his**, and **hers**

* Edit distance between “he” and “she” = 1

‘s’ should be inserted in “he” in order to turn into “she”

* Edit distance between “he” and “his” = 3

‘e’ should be deleted in “is” should be inserted.

* Edit distance between “he” and “hers” = 2

‘r’ and ‘s’ should be inserted

* Edit distance between “she” and “his” = 4

‘s’ and ‘e’ should be deleted and ‘i’ and ‘s’ should be inserted

* Edit distance between “she” and “hers” = 3

‘s’ should be deleted. ‘r’ and ‘s’ should be inserted

* Edit distance between “his” and “hers” = 3

‘i’ should be deleted and ‘e’, ‘r’ should be inserted.