Mid Term Examination

CS 513 KDD

Name: Manan Bhatt

CWID: 10453036

1. The given distant function is correct, because in the distant function follows all the properties that should be true for any distant function

A (0,0,0), B (0,1,0), C (0,1,1), D (1,1,1)

The Distance between A and B: D (A, B) = ((0-0)2+(1-0)2 +(0-0)2

= 1

The Distance between B and C: D (B, C) = (0-0)2+(1-1)2 +(0-1)2

= 1

The Distance between C and D: D (C, D) = ((0-1)2+(1-1)2 +(1-1)2

= 1

The Distance between A and D: D (A, D) = (0-1)2+(0-1)2 +(0-1)2

= 3

The Distance between A and C: D (A, C) = (0-0)2+(0-1)2 +(0-1)2

= 2

The Distance between B and D: D (B, D) = (0-1)2+(1-1)2 +(0-1)2

= 2

This equation follows the three properties that should be true for distance function i.e

1. **D(x,y) >= 0 and D(x,y) = 0 if x=y**

For our values distance between any two points is greater than zero.

1. **D(x,y) = D(y,x)**

As we are taking the square of distance the value for D(x,y) will always be equal to D(y,x)

1. **D(x,z)<= D(x,y) + D(y,z)**

For our 4 points summation of distance of any two points is always greater than other point which satisfies this property as well.

Q8

1. Estimate the number of cases
2. USA:

Population: 331 Million

Cases per Million: 381.24

Total Number of Cases in USA = 331\*381.24

= 1,26,190.44

1. Italy:

Population: 60 Million

Cases per Million: 1463.97

Total Number of Cases in Italy = 60\*1463.97

= 87,838.2

1. Spain:

Population: 47 Million

Cases per Million: 1590.24

Total Number of Cases in Italy = 47\*1590.24

= 74,741.28

1. Given the person is living in USA what is the probability of being infected

P (Infected / Living = USA) = 1,26,190.44 / 331 Million

= 0.0003

1. Given the person is infected what is the probability of being in USA

P (USA / Infected = Y) = 1,26,190.44 / (1,26,190.44 + 87,838.2 + 74,741.28)

= 1,26,190.44 / 2,87,769.92

= 0.4385