2018AIML600

# Ans 1.

Sorted Data: = { 1, 7, 8, 11, 12, 16, 22, 25, 30, 32, 37, 37, 42, 57, 61, 68 }

Number of bins = 4.

## Equal Frequency Binning

freq. = 16/4 = 4.

Bin1 : { 1, 7, 8, 11 }

Bin2 : { 12, 16, 22, 25 }

Bin3 : { 30, 32, 37, 37 }

Bin4 : { 42, 57, 61, 68 }

## Equi-Width Binning

Bin\_width = (max – min)/no of bins = (68 – 1)/4 = 16.75 = 17.

Bin1 : [1,18) => { 1, 7, 8, 11, 12, 16 }

Bin2 : [18, 35) => { 22, 25, 30, 32 }

Bin3 : [35, 52) => { 37, 37, 42 }

Bin4 : [52, 69) => { 57, 61, 68 }

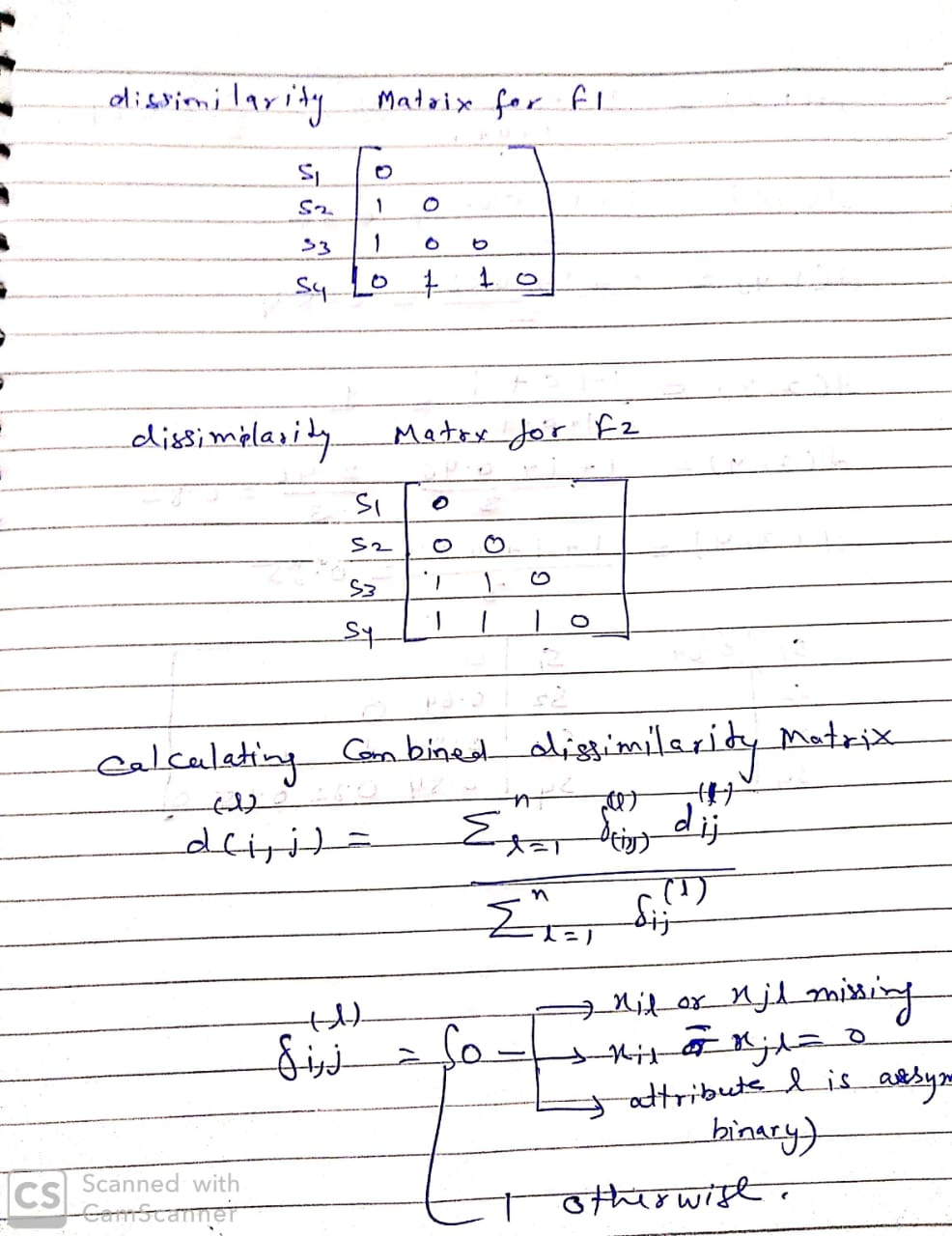
We can see that equal frequency binning contains all bins with equal number of elements (exception with last bin in some special cases).The drawback of this types of binning is that same elements may occur in more than one bins if appeared more than once in actual datasets.

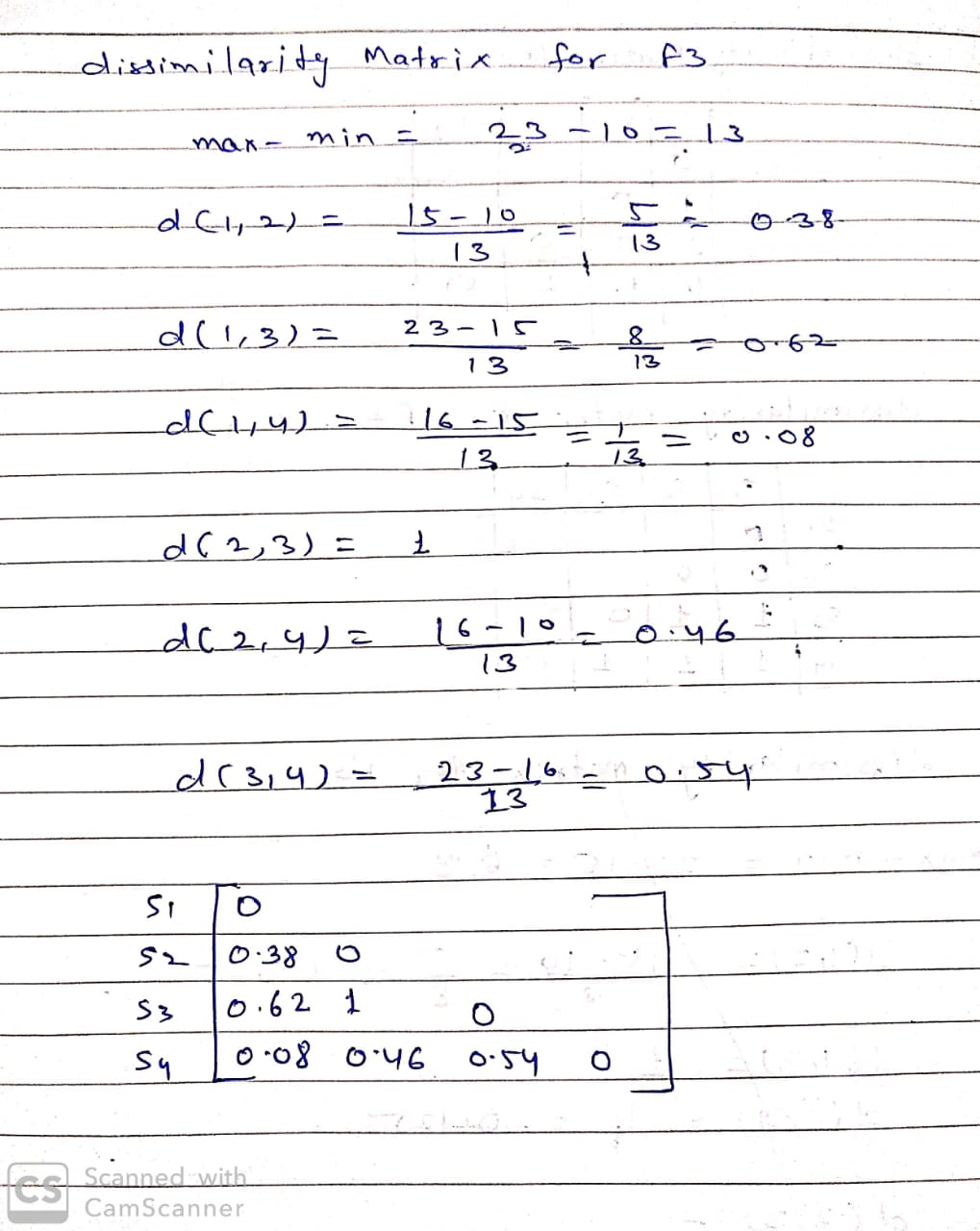
Equi-Width Binning, the width of bin is same for all, not the number of elements. No data duplicity is possible across the different bins in this types of bins.

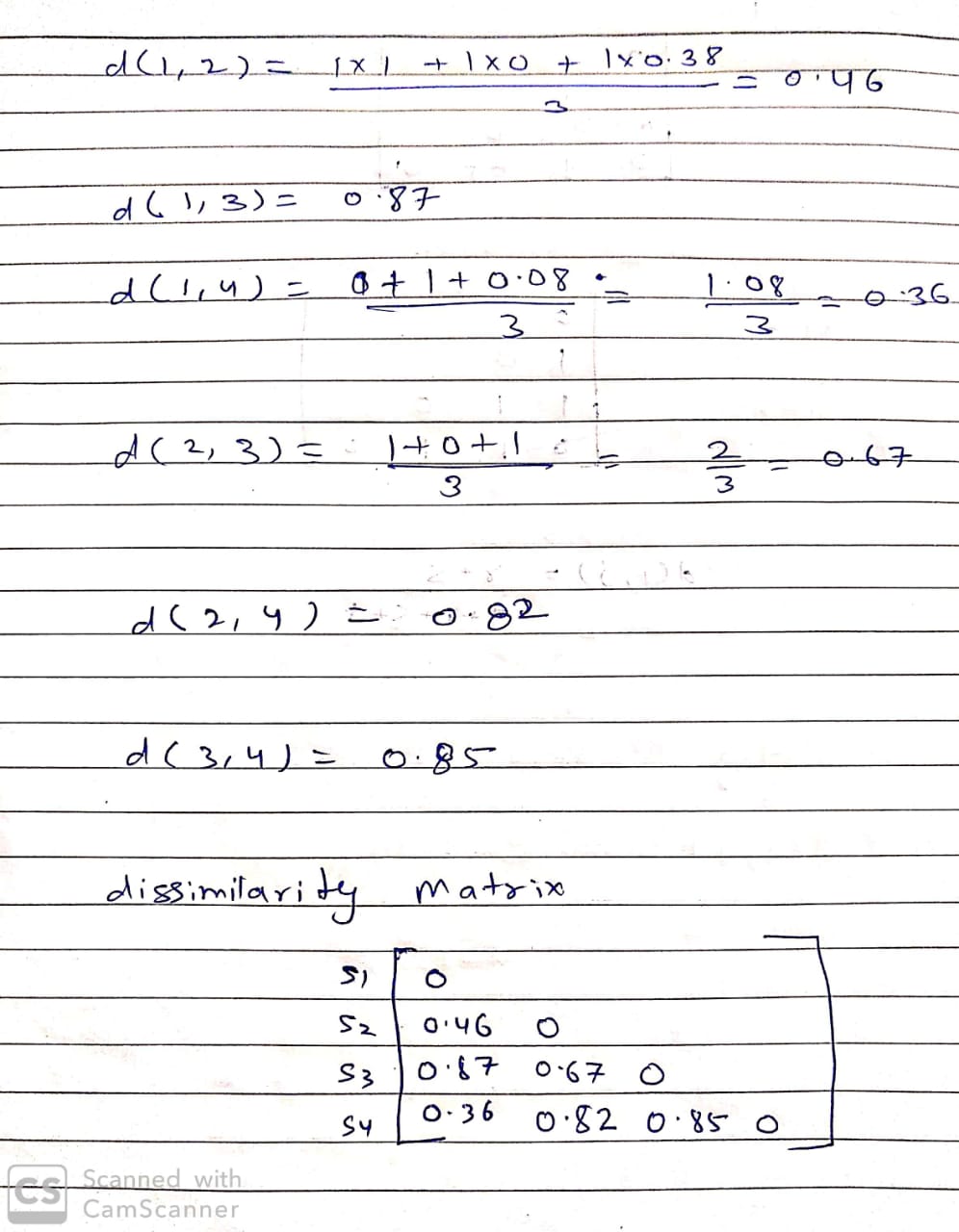
Ans 2.

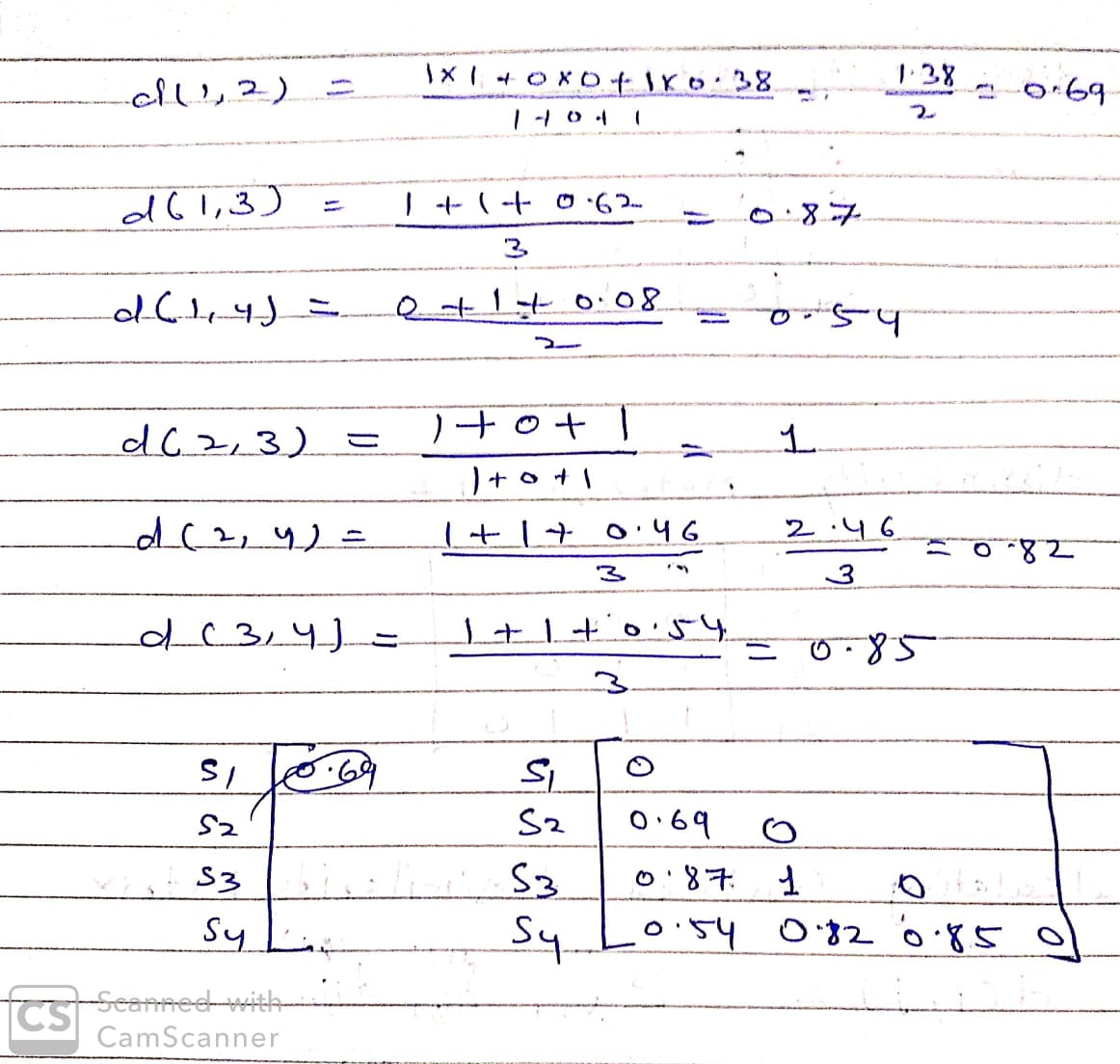
(a)

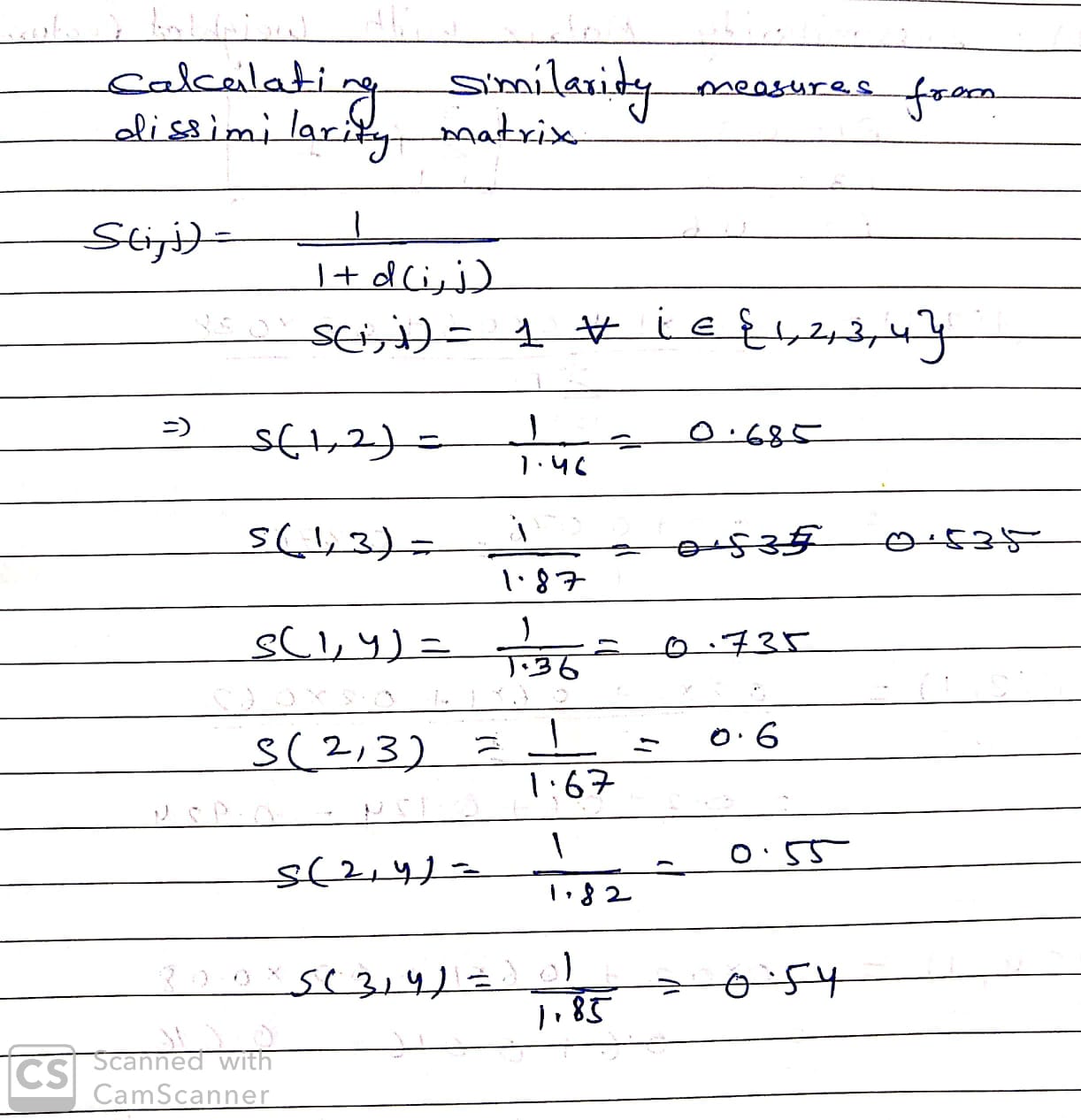
Solution is given as below

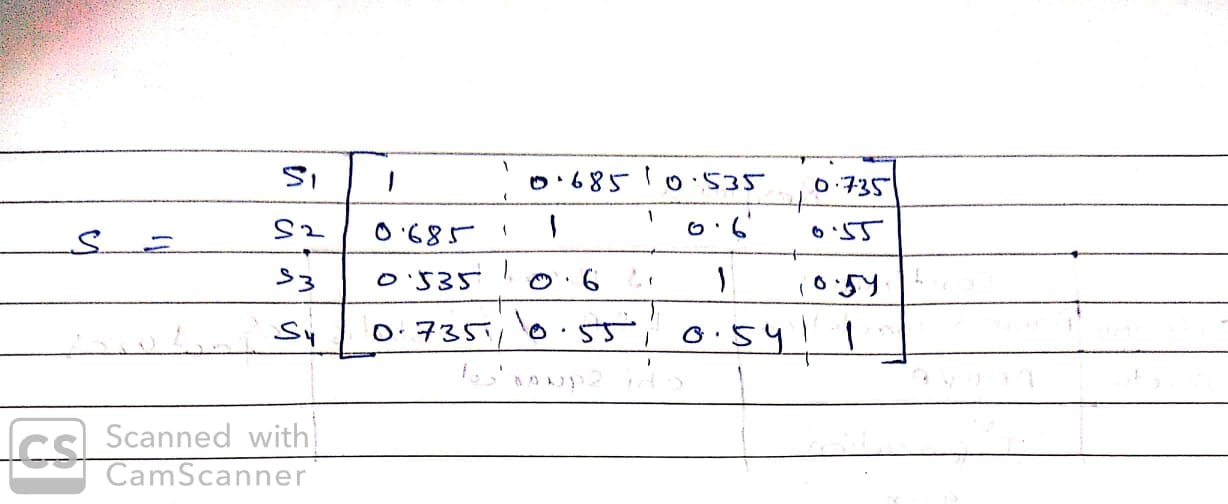






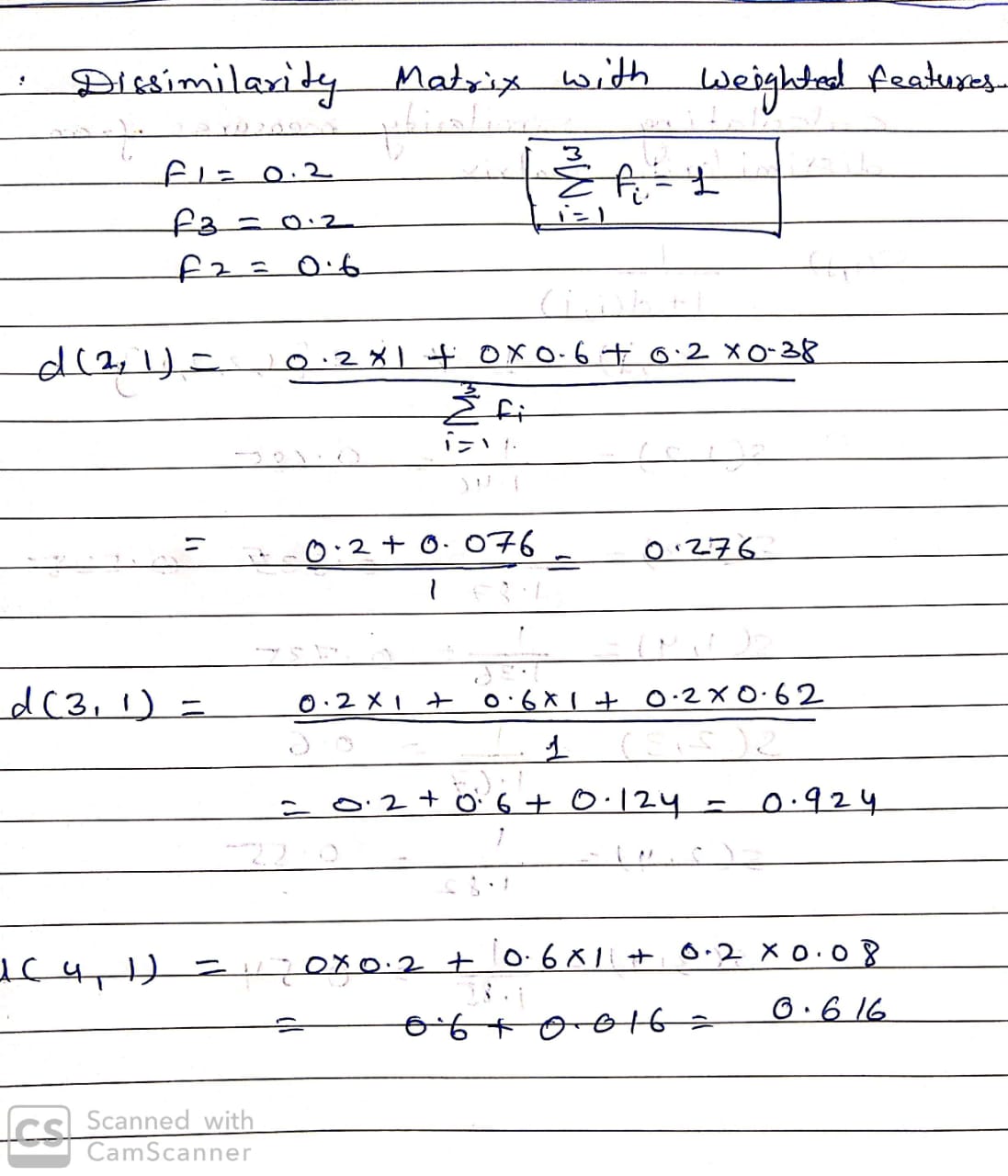


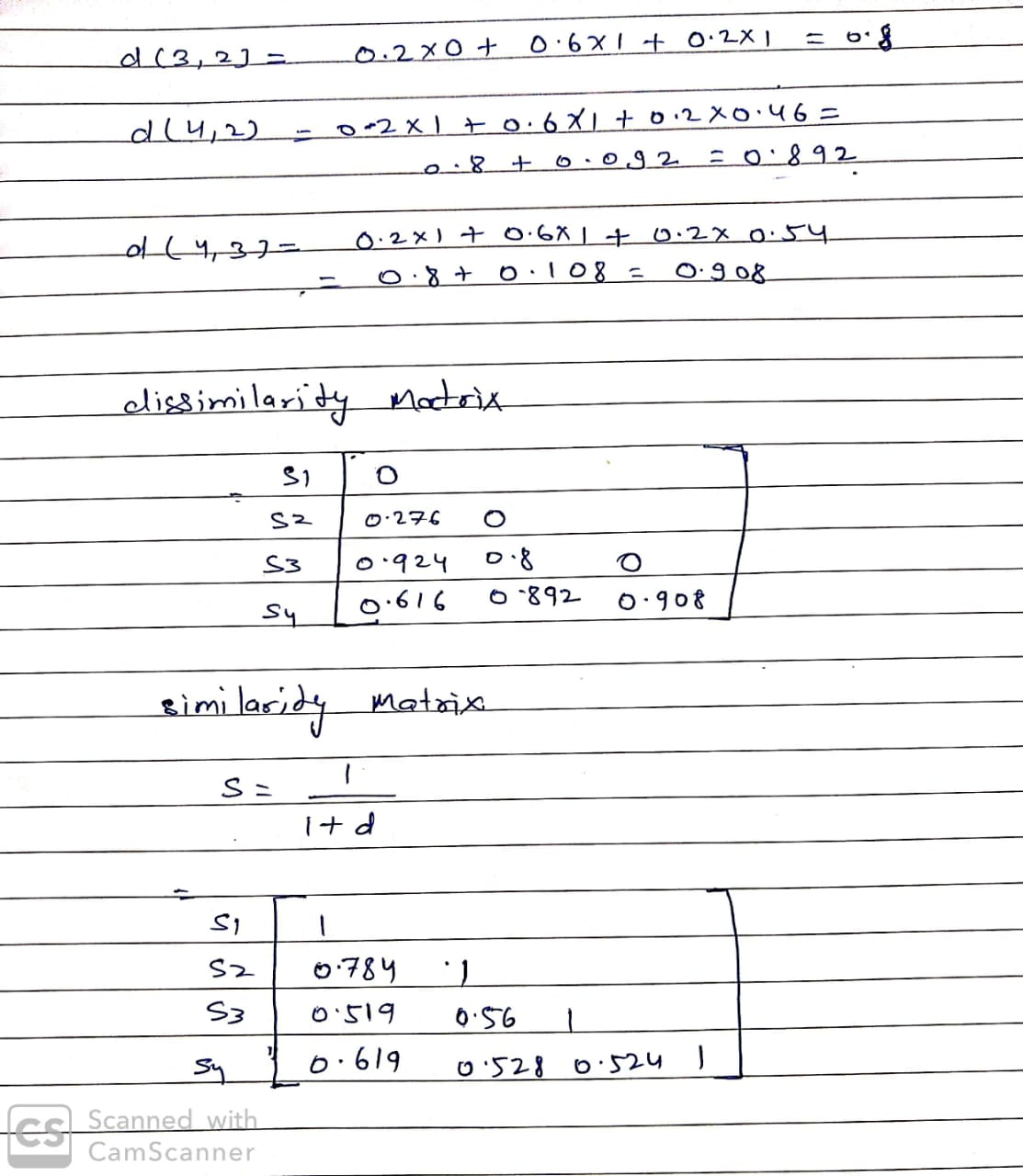




(b)

Using the dissimilarity matrices calculated in part (a)





Ans 3

Find the below

