NAME : Muhammad Younis Afridi

Course : BS(SE)

ID : 12701

Subject : Data Mining.

Constructor : Sir Zain Shoukat.

Q1 Implement a Code of Genetic Algorithm in any language and show the out - put ?

Ans

```
Class Genetic :
    def __init__ (self, string) :
        Self. population = []
        for i in rang (100) :
            Sample = []
        for i in range (len (string)) :
            # 97 a   122 z
            Sample. append (chr (int (random (97, 123))))
    Self. population. append (Sample)
    Self. mutation_ rate = 0.1
    def fitness (self, Sample) :
        Count = 0
        for i, j in Zip (Sample, self. String) :
            if i == j :
                Count + = 1
        return Count.
```

P.T.O

```
def Selection (self) :
    self . population . sort (key = lambda x: self fitness(x),
                              reverse = True )
    new_ population = self . population [:40]
    samples = rn. sample (self . population [40:], 10)
    for i in sample :
        new_ population . append (i)
    self . populaton = new_ population
def crossover (self)
    for i in range ( 100 - len (self . population) ):
    parent = rn . sample (self . population, 2)
    new_ child = ''
    for i in range ( len self (string)) :
        if random (0, 1) < 0.5 :
            new_ child + = parent [0] [i]
    else :
            new_ child + = parend [1] [i] ,
    self . population . append (new_ child] .
```

```
def mutation (self) :
    new_population = []
    for sample in self.population :
        mutated = ''
        for i in sample :
            if random (0,1) < self.mutation_rate :
mutated += chr (int (random (97,123)))
else :
    mutated+ = i
new_population. append (mutated)

Self. population = new_population

while True :

    generation + = 1
print (generation, self.fitness( self.population [0]))

if self.fitness (self.population [0] == len (self.string)

    print (self.population [0])
    return self. population [0]
~~else :~~
       ~~return none~~
elif generation > 10000 :
    return none.
self . selective ()
self . crossover ()
self . mutation ()

g = Genetic (hellower ld')
```

Q2 Implement a code of fuzzy logic in an language and show the out put ;

Ans :

```
# include <stdio.h>
# include <conio.h>
# include <math.h>
float min (float a, float b);
float max (float a, float b);
int main ()
{
    float fa, fb, fi, fu, fac;
    float x = 1.0;
print ("enter membership function of first set:");
Scanf ("%f", and fb);
fi = min (fa, fb);
fu = max (fa, fb);
fac = x - fai;
printf(" The membership function of intersection = x0.%f\n",fi
printf( The membership function of union = x0.%f\n", fu);
```

```
printf( The membership function of Complement of first Set =x0 of %f

    return 0 ;
  float  min (float a , float b)
{
      if (a<b)
        return a ;
      else
        return b ;
  }

      float max (float a, float b)
        if  (a>b)
          {
             return a ;
      }
         else
      {
         return b ;
      }
```

Q3 Solve this using KNN .

	Name	Acid durability	Strength	Class
i	Type_1	7	7	Bad
ii	Type-2	7	4	Bad
iii	Type-3	3	4	Good
iv	Type-4	1	4	Good

Test-Data → acid durability = 3, and strength = 7

Class = ?

Ans 3 $K = 3$ nearest neighbours

So

Euclidean distance $= \sqrt{(x_H + H_1)^2 + (x_w - W_1)^2}$

X_H = observed value

H_1 = Actual value

$D(x, i) = \sqrt{(3-7)^2 + (7-7)^2}$

$= 4$

$D(x, ii) = \sqrt{(3-7)^2 + (7-4)^2}$

$= \sqrt{16+9} = \sqrt{25}$

$= 5$

$D(x, iii) = \sqrt{(3-3)^2 + (7-4)^2}$

$= 3$

$D(x, iv) = \sqrt{(3-1)^2 + (7-4)^2}$

$= \sqrt{4+9}$

$= 3.6$

N_1 = nearest value.

So values are

$D(x, i) = 4 \rightarrow N_3 \rightarrow$ Bad.

$D(x, ii) = 5$

$D(x, iii) = 3 \rightarrow N_1 \rightarrow$ good

$D(x, iv) = 3.6 \rightarrow N_2 \rightarrow$ good

Therefore

acid durability = 3 and Strength = 7 will

belong to class True.

Ans

Q4 Give solved examples of hierarchical clustering.

Ans

	X	Y
Z_1	0.4	0.53
Z_2	0.22	0.38
Z_3	0.35	0.32
Z_4	0.26	0.19
Z_5	0.08	0.41
Z_6	0.45	0.30

Step 2

Find the distance matrix by using manhattan distance which is given by distance between the points $(P_1, P_2) + (q_1, q_2)$ is given by $|P_1 - q_1| + |P_2 - q_2|$

Distance between the points

$$Z_1, Z_2 = |0.4 - 0.22| + |0.22 - 0.38|$$

$$= |0.18| + |0.15|$$

$$= 0.33$$

Distance (Z_1, Z_2)

Distance (Z_1, Z_3) $= |0.4 - 0.35| + |0.53 - 0.32|$

$= |0.05| + |0.21|$

$= 0.26$

Distance (Z_1, Z_4) $= |0.4 - 0.26| + |0.53 - 0.32|$

$= |0.14| + |0.21|$

$= 0.35$

Distance (Z_1, Z_5) $= |0.4 - 0.08| + |0.53 - 0.41|$

$= |0.32| + |0.12|$

$= 0.44$

Distance (Z_1, Z_6) $= |0.4 - 0.45| + |0.53 - 0.30|$

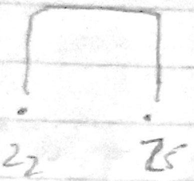
$= 0.28$

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Z_1	0	0.33	0.24	0.35	0.44	0.28
Z_2	0.33	0	0.19	0.23	0.01	0.31
Z_3	0.26	0.19	0	0.23	0.36	0.12
Z_4	0.35	0.23	0.22	0	0.4	0.3
Z_5	0.44	$\boxed{0.01}$	0.36	0.4	0	0.48
Z_6	0.28	0.31	0.12	0.3	0.48	0

- The minimum distance is between

P.T.O

The points Z_2 + Z_5 and they both from the first cluster.



The dendrogram is drawn to represent the first cluster formatting.

<u>Step 3</u>: update the distance matrix using the first cluster (Z_2, Z_5).

There are three methods to do this

1. Single link.
2. Complex link.
3. Average link.

distance between (Z_2, Z_5) and Z_1 is given by using single link ⇒

$$\text{Min}\{(Z_2, Z_1), (Z_5, Z_1)\}$$

$$= \text{Min}(0.32, 0.44)$$

$$= 0.33$$

Distance between (Z_2, Z_5) and Z_1 using Complete link is given by.

(11)

Max $[(z_2, z_1), (z_5, z_1)]$.

$= \text{Max} \{0.33, 0.44\}$.

$= 0.44$

dist $[(z_2, z_5) + z_1)$ using Average link

given by $= \frac{1}{2} [\text{dist}(z_2, z_1) + \text{dist}(z_5, z_1)]$.

$= \frac{1}{2} [0.33 + 0.44]$.

$= 0.55$

The update matrix is given by.

	z_1	$z_2 z_5$	z_3	z_4	z_6
z_1	0	0.33	0.26	0.35	0.28
z_2, z_5	0.32	0	0.19	0.23	0.31
z_3	0.26	0.19	0	0.22	$\boxed{0.12}$
z_4	0.35	0.23	0.22	0	0.3
z_6	0.28	0.31	$\boxed{0.12}$	0.3	0

distance between $(z_2, z_5) + (z_3) = 0.19$

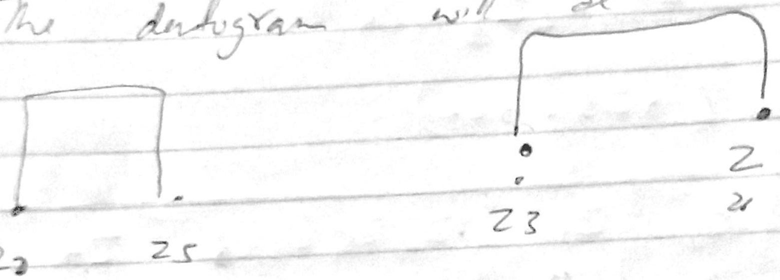
dist $[(z_2, z_5) + (z_4)] = 0.23$

P.T.O

dist $[(Z_2, Z_5) + (Z_6)] = 0.31$

* The minimum distance is between $Z_3 + Z_6$ and they form the second cluster.

The dendogram will be



Step 4

Again update the distance matrix using the newly form clustered

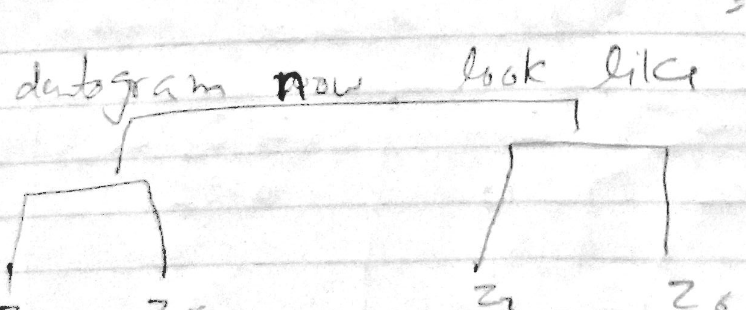
dist $((Z_3 Z_6), Z_4)$
$= \min(0.22, 0.3)$
$= 0.22$

	Z_1	Z_2, Z_5	$Z_3\ Z_6$	Z_4
Z_1	0	0.33	0.26	0.31
$Z_2 Z_5$	0.33	0	0.19	0.23
$Z_3 Z_6$	0.26	0.19	0	0.22
Z_4	0.35	0.23	0.22	0

distance $(Z_2, Z_5) + (Z_3 Z_6)$
$= \min(0.19, 0.31)$
$= 0.19$

dist $((Z_3 Z_6), Z_1)$
$= \min(0.26, 0.28)$
$= 0.26$

The dendogram now look like



Step 5 The update distance matrix is again

min (Z_1, Z_2Z_3, Z_4)

min (0.33, 0.26)

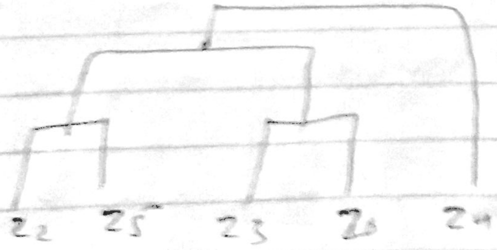
= 0.3326

distance (Z_4, $Z_2Z_5Z_3Z_6$)

min (0.23 . 0.22)

Dist	Z_1	$Z_2Z_5Z_3Z_6$	Z_4
Z_1	0	0.26	0.36
$Z_2Z_5Z_3Z_6$	0.26	0	0.22
Z_4	0.36	0.22	0

Now dendogram looks.



Step 6 update distance matrix

	Z_1	$Z_2Z_5Z_3Z_6Z_4$
Z_1	0	0.33
$Z_2Z_5Z_3Z_6Z_4$	0.33	0

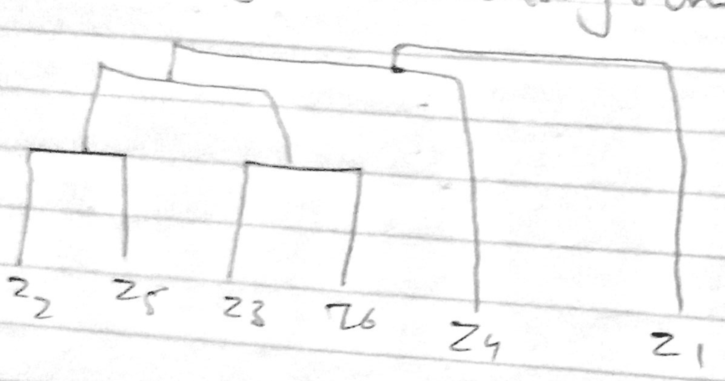
(14)

distance betwe $(z_1 + (z_2 z_5 z_3 z_6 z_4))$

$= (\min(0.33, 0.35))$

$= 0.33.$

The final dendrogram is represented



END