Name : Shubham Dutta

Year : 2nd

Section : 2B

Stream : CST

Roll No. :- 58

Enrollment No. : 12019009022112

Paper Name : Artificial Intelligence & Machine Learning Advanced

Paper Code : PCC-CS405

Signature : Shubham Dutta

Date : 11/05/2021

Answers

1. A) N1 : Let N1 be 10 (between 10 & 99)

∴ N2 : 10 * 10 = 100

∴ N3 : 10 * 100 + 100 = 1100

∴ N4 : 12

∴ N5 : 12 * 10 * 100 + 1100 = 13100

∴ N6 : 112

∴ N7 : 1100 * 100 = 110000

Decimal scaling of the 7 numbers

Max value : 110000

∴ j = 6 ∴ $(v^i / 10^j)$.

i.e. N / 1000000

After Normalisation :-

N1 : 10/1000000 = 1×10^{-5}

N2 : 1×10^{-1}

N3 : 1.1×10^{-3}

N4 : 1.2×10^{-5}

N5 : 0.0131

N6 : 1.12×10^{-1}

N7 : 0.11

∴ Z-score of following numbers :-

S.D = 40448.88 , Mean = 17776.28

∴ After Normalisation :-

N1 = -0.433 , N2 = -0.431 ,

N3 = -0.407 , N4 = -0.433 ,

N5 = -0.114 , N6 = -0.431

N7 = 2.25

---

2. B)

X	Y	x^2	y^2	xy
90	80	8100	6400	72000
75	85	5625	7225	6875
70	60	4200	3600	4200
60	55	3600	3025	3300
50	65	2500	4225	3250

Spearman form

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$= 24725 - \frac{(345)^2}{5}$$

$$= 920$$

$$S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n}$$

$$= 24325 - \frac{345 \times 345}{5}$$

$$= 520$$

$$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{5}$$

$$= 24975 - \frac{345^2}{5}$$

$$= 670$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{520}{\sqrt{670 \times 920}} = 0.662$$

3.B)

	P1	P2	P3	P4	P5	P6
P1	0					
P2	0.23	0				
P3	0.22	0.15	0			
P4	0.37	0.20	0.15	0		
P5	0.34	0.14	0.28	0.29	0	
P6	0.23	0.25	0.11	0.22	0.39	0.

To update the distance Matrix MIN.

MIN [dist (P5, P6, P1)]

⇒ MIN [dist (P3, P1), (P6, 1)]

= min (0.22, 0.23)

= 0.22

To update the distance matrix MIN [dist (P3, P6), P2]

MIN (dist (P3, P2), (P6, P2))

= min (0.15, 0.25)

= 0.15

To update the distance matrix MIN [dist (P3

MIN [[dist (P3, P6), P4)]

⇒ MIN [dist (P3, P4), (P6, P4)]

= min (0.15, 0.22)

= 0.15

To update the distance matrix.

MIN [(dist (P2, P5), (P3, P6)), P4]

= min ((0.20, 0.15)

= 0.15

To update the distance Matrix.

MIN [dist (P2, P5, P3, P6), P1]

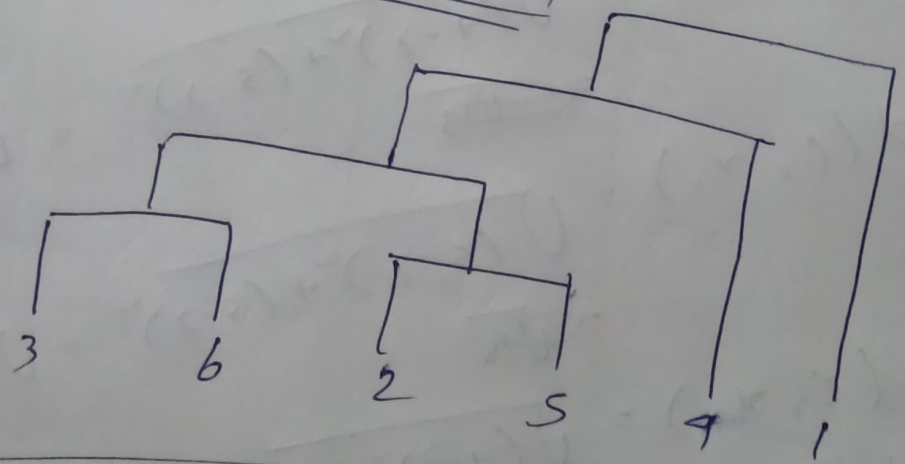
= min (0.22, 0.37)

= 0.22

Update distance matrix for cluster. P2, P5, P3, P6, P4

	P1	P2, P5, P3, P6 P4
P1	0	
P2, P3, P4, P6, P5	0.22	0

DENDOGRAM DIAGRAM



4 b) $C_1 = \{x_1, x_3\}$; $C_2 = \{x_2, x_4\}$

So, $C_1 = \{\frac{2+6}{2}, \frac{2+8}{2}\} = \{4, 5\}$

$C_2 = \{\frac{8+2}{2}, \frac{6+4}{2}\} = \{5, 5\}$

Iteration - I

	C_1	C_2	Membership
x_1	$\sqrt{13}$	$\sqrt{18}$	C_1
x_2	$\sqrt{17}$	$\sqrt{10}$	C_2
x_3	$\sqrt{13}$	$\sqrt{10}$	C_2
x_4	$\sqrt{5}$	$\sqrt{10}$	C_1

$d(C_1, x_1) = \sqrt{(4-2)^2 + (5-2)^2} = \sqrt{13}$

$d(C_2, x_1) = \sqrt{(5-2)^2 + (5-2)^2} = \sqrt{18}$

$d(C_1, x_2) = \sqrt{(4-8)^2 + (5-6)^2} = \sqrt{17}$

$d(C_2, x_2) = \sqrt{(5-8)^2 + (5-6)^2} = \sqrt{10}$

$d(C_1, x_3) = \sqrt{(4-6)^2 + (5-8)^2} = \sqrt{13}$

$d(C_2, x_3) = \sqrt{(5-6)^2 + (5-8)^2} = \sqrt{10}$

$$d(c_1, x_4) = \sqrt{(4-2)^2 + (5-4)^2} = \sqrt{5}$$

$$d(c_2, x_4) = \sqrt{(5-2)^2 + (5-4)^2} = \sqrt{10}$$

Iteration - II

$$C_1 = \{x_1, x_4\}, \quad C_2 = \{x_2, x_3\}$$

$$C_1 = \left\{\frac{2+6}{2}, \frac{2+8}{2}\right\} = \{4, 5\}$$

$$C_2 = \left\{\frac{8+6}{2}, \frac{6+8}{2}\right\} = \{7, 7\}$$

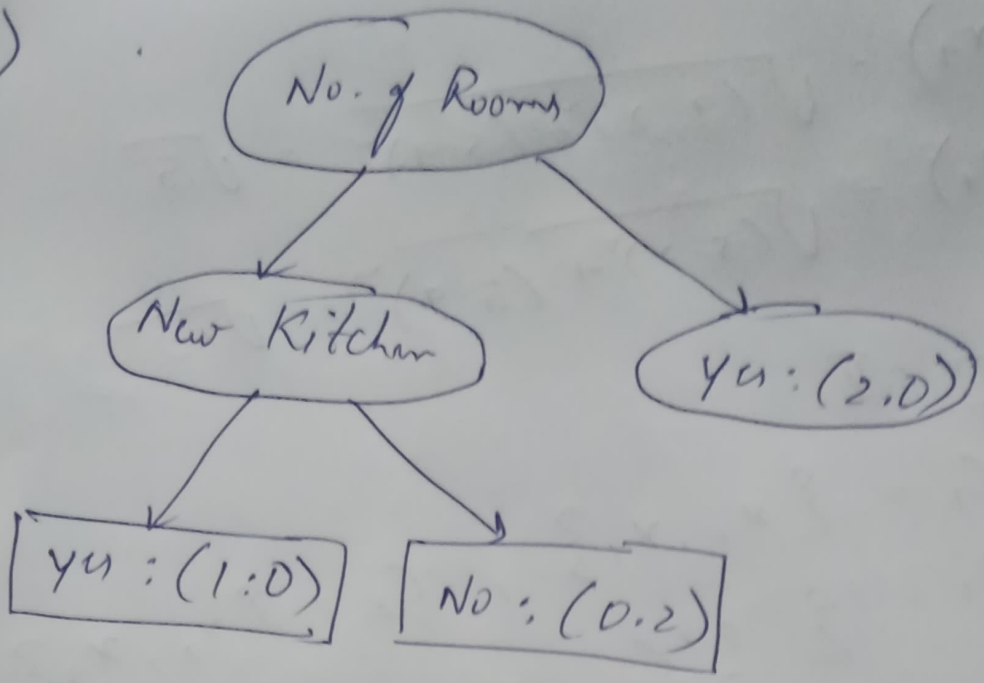
5. A)

House	Furniture	No. of Rooms	New Kitchen	Acceptable
1	No	3	Yes	Yes
2	Yes	3	No	No
3	No	4	No	Yes
4	No	3	No	No
5	Yes	4	No	Yes

a) $H(S, \text{Acceptable}) = -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)$

$= 0.971$

6)



7. B) CNN - Architecture.

In Deep learning, a convolution Neural Network (CNN) is a class of deep neural network, most commonly applied to analysing visual imagery. CNN are state of art models for Image.

Classification segmentation: Object Detection & many other image processing tasks: In order to get started with the field of Image processing or to improve the ~~production~~ prediction accuracy of custom CNN models.

Following Architectures :-

- VGG Net
- Res Net
- Dense Net
- Inception Net

---

6. A) Random Forest is a flexible, easy to use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms, because of its simplicity & diversity (it can be used for both classification & regression tasks)

Random Forest would clockwise with 2 trees