University of Engineering & Management, Kolkata

Even Semester Term- II Examination, May, 2021

Course: B.Tech(CS) **Semester:4th**

Paper Name: Artificial Intelligence & Machine Learning Advanced

Paper Code: PCC-CS405

Full Marks: 70 **Time: 2 hours**

Answer all questions. Each question is of 10 marks.

1. **A)** Normalize the following 7 numbers using Decimal-Scaling and Z-Score normalization. The numbers are as follows:
N1: A random number between 10 to 98
N2: 10*N1
N3: 10*N2 + N2
N4: 12
N5: N4*N1*100 + N3
N6: N4+100
N7: N3*100

OR

B) Normalize the following 7 numbers using Decimal-Scaling and Z-Score normalization. The numbers are as follows:
N1: A random number between 10 to 58
N2: 10*N1
N3: 10*N2 + N2
N4: 50
N5: N4*N1*100 + N3
N6: N4+100
N7: N6*100

2. A) From the table below established the relationship between X and Y Established the relationship between X and Y using least squared linear regression. Find the R squared error of estimate. 5+5

OR

B) From the table below find the Pearson's Coefficient of Correlation. Established the relationship between X and Y using least squared linear regression. 5+5

X	Y
90	80
75	85
70	60

60	55
50	65

3. A) The given dataset is {E,A,C,B,D}; The distance matrix is given below.

	E	A	C	B	D
E	0				
A	1	0			
C	2	2	0		
B	2	5	1	0	
D	3	3	6	3	0

Find the final clusters using hierarchical agglomerative clustering. Also draw the Dendrogram diagram.

OR

B) The given dataset is {P1,P2,P3,P4,P5,P6}; The distance matrix is given below.

	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

Find the final clusters using hierarchical agglomerative clustering. Also draw the Dendrogram diagram.

4. A) For the below dataset find out the final clusters using K-Means algorithm. Note: i) K=2; ii) Use Manhattan distance

X	Y
1	1
1.5	2
3	4
5	7

3.5	5
4.5	5
3.5	4.5

OR

B) We are given the following four data points in two dimension: XI = (2,2), X2 =(8,6), X3 =(6,8), X4 = (2,4). We want to cluster the data points into two clusters CI and C2 using the K-Means algorithm. Euclidean distance is used for clustering. To initialize the algorithm we consider CI ={xI,x3} and C2 ={x2,x4}. After two iteration of the K-means algorithm, find the cluster memberships

5. A) Consider the dataset in the given table, of houses by five training examples. The target attribute is 'Acceptable', which can have values 'yes' or 'no'. This is to be predicted based on the other attributes of the house.

House	Furniture	Number 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) Compute entropy of target attribute.
b) Construct a decision tree from the above examples, that would be learned by the ID3 algorithm.

OR

B) Consider the dataset in the given table, of houses by five training examples. The target attribute is 'Acceptable', which can have values 'yes' or 'no'. This is to be predicted based on the other attributes of the house.

House	Furniture	Number 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

Construct a decision tree from the above examples, that would be learned by the CART algorithm.

6. A) Discuss the random forest building algorithms with an example.

OR

B) Consider the following data set and predict the class using KNN for k=5 : age<=30, income=medium, student=yes, institute status=fair.

Data	Age	Income	Student	Institute status	Education loan
D1	<=30	High	No	Fair	No
D2	<=30	High	No	Excellent	No
D3	30…40	High	No	Fair	Yes
D4	>40	Medium	No	Fair	Yes
D5	>40	Low	Yes	Fair	Yes
D6	>40	Low	Yes	Excellent	No

D7	31…40	Low	Yes	Excellent	Yes
D8	<=30	Medium	No	Fair	No
D9	<=30	Low	Yes	Fair	Yes
D10	>40	Medium	Yes	Fair	Yes
D11	<=30	Medium	Yes	Excellent	Yes
D12	31…40	Medium	No	Excellent	Yes
D13	31…40	High	Yes	Fair	Yes
D14	>40	medium	no	Excellent	No

7. A) Demonstrate how the perceptron model can be used to represent the AND and OR functions between a pair of Boolean variables.

OR

B) Discuss CNN architecture.