

Time: 3 hours

Max. Marks: 80

N.B. (1) Question one is Compulsory.

(2) Attempt any 3 questions out of the remaining.

(3) Assume suitable data if required.

Q. 1 a) Demonstrate with diagram Data Mining Architecture. 05

b) What is outlier? Explain different types of outliers 05

c) What is prediction? Explain about Linear regression method. 05

d) Explain different OLAP operations on multidimensional data . 05

Q. 2 a) Briefly outline with example, how to compute dissimilarity between the objects describe following

i) Nominal attributes

ii) Asymmetric binary attributes 10

b) Calculate Accuracy, Recall and Precision with the help of following data:

True Positive (TP)= 45, True Negative (TN) = 25, False Positive (FP)= 18, False

Negative (FN)= 12. 10

Q. 3 a) Describe the various phases in knowledge discovery process with a neat diagram 10

b) What is web mining? Explain web content mining in detail 10

Q. 4 a) What is Clustering Technique ? Discuss the Agglomerative algorithm with the following data and plot a Dendrogram using single link approach. The table below comprises sample data items indicating the distance between the elements. 10

Item	P1	P2	P3	P4	P5
P1	0				
P2	9	0			
P3	3	7	0		
P4	6	5	9	0	
P5	11	10	2	8	0

b) A database has five transactions. Let min-support = 60_ and min- confidence = 80%. Find all frequent item sets by using Apriori Algorithm T_ID is the transaction ID 10

T_ID	Items bought
T1000	M,O,N,K,E,Y
T1001	D,O,N,K,E,Y
T1002	M,A,K,E
T1003	M,U,C,K,Y
T1004	C,O,O,K,E

- Q. 5 a) Explain Decision Tree classification Approach in detail ? 10
b) What is market basket analysis? Explain with an example. State and explain with formula the meaning of the following term:
i) Support
ii) Confidence 10
- Q. 6 a) Explain BIRCH algorithm using appropriate diagram. 10
b) Define multidimensional and multilevel association mining 10
-