CMR GROUP OF INSTITUTIONS
EXPLORE TO INVENT

CMR INSTITUTE OF TECHNOLOGY
UGC AUTONOMOUS
Approved by AICTE, New Delhi & Permanently Affiliated to JNTUH, Hyderabad & Accredited by NBA

DEPARTMENT OF CSE (AI/ML) & (DS)

Data Mining and Data Analytics (20-CS-PC-312)

Mid – 1 Assignment – Questions

III – B.TECH I – SEM –CSE SUBMISSION DEADLINE: 12-09-2022

Note: SET-1 writes first 25% students
SET-2 writes next 25% students
SET-3 writes next 25% students
SET-4 writes last 25% students

ANSWER ALL THE QUESTIONS

SET-1

S.No	Question	BTL	CO	PO	Unit
1	Define Data and Information. Explain kinds of Data with an example	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
2	Categorize the Data mining functionalities in detail.	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
3	Elaborate the different data preprocessing techniques 1) Apply the following rules on a database has five transactions. Let min sup =60% and min items bought conf=80% TID Items bought T100 {M,O,N,K,E,Y} T200 {D,O,N,K,E,Y} T300 {M,A,K,E} T400 {M,U,C,K,Y} T500 {C,O,O,K,I,E} a) Find all frequent item sets using Apriori b) List all of the strong association rules (with supports and Confidence)	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
4	What is FP-Growth tree? Explain FP-Growth Tree Algorithm with an Example	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
5	Explain Decision tree induction algorithm for classification.Discuss the usage of information gain.	VI	CO 3	PO 1, PO 2, PO 3, PO 6	III

SET-2

S.No	Question	BTL	CO	PO	Unit
1	What is Data Mining? Explain Data Mining preprocessing Techniques.	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
2	List out Data mining task primitives	VI	CO 1	PO 1, PO 2, PO 3, PO 6	I

3	Illustrate with an A-priori algorithm for the given dataset below. TID List of items 001 milk, dal, sugar, bread 002 Dal, sugar, wheat,jam 003 Milk, bread, curd, paneer 004 Wheat, paneer, dal, sugar 005 Milk, paneer, bread 006 Wheat, dal, paneer, bread	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
4	Apply the following rules on a database has five transactions. Let min sup =60% and min items bought conf=80% TID Items bought T100 {M,O,N,K,E,Y} T200 {D,O,N,K,E,Y} T300 {M,A,K,E} T400 {M,U,C,K,Y} T500 {C,O,O,K,I,E} Apply FP-Growth algorithm to the following transactional data to find frequent itemsets. List all frequent itemsets with their support count.	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
5	Explain naïve Bayes algorithm with an example.	VI	CO 3	PO 1, PO 2, PO 3, PO 6	III

SET-3

S.No	Question	BTL	CO	PO	Unit
1	What is KDD? Explain KDD steps with a neat diagram.	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
2	List out issues in Data Mining and Data mining task primitives	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
3	Explain improving the efficiency of Apriori Algorithm	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
4	Apply FP-Growth algorithm to the following transactional data to find frequent itemsets. List all frequent itemsets with their support count. TID List of Item IDs 1 I1, i3,i5,i7 2 I2, i4,i6,i8 3 I1, i3,i5,i7 4 I9, i7,i5,i1 5 I2, i4,i6,i7 6 I1, i2,i3,i4 7 I3, i4,i5,i6 8 I7, i8,i6,i1 9 I8, i5,i3,i2 10 I9, i3,i4,i6	VI	CO 2	PO 1, PO 2, PO 3, PO 6	II
5	Explain naïve Bayes algorithm with an example.	II	CO 3	PO 1, PO 2, PO 3, PO 6	III

SET-4

S.No	Question	BTL	CO	PO	Unit
1	What is KDD? Explain KDD steps with a neat diagram	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
2	Write a short notes on a)Data Cleaning b) Data Transformation	II	CO 1	PO 1, PO 2, PO 3, PO 6	I
3	Illustrate with an A-priori algorithm for the given dataset below. TID List of items 001 milk, dal, sugar, bread 002 Dal, sugar, wheat,jam 003 Milk, bread, curd, paneer 004 Wheat, paneer, dal, sugar 005 Milk, paneer, bread 006 Wheat, dal, paneer, bread	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
4	What is Correlation Analysis? Explain correlation Analysis from Association Analysis.	II	CO 2	PO 1, PO 2, PO 3, PO 6	II
5	Explain Decision tree induction algorithm for classification.Discuss the usage of information gain.	II	CO 3	PO 1, PO 2, PO 3, PO 6	III

****