

F.Y. MCA - Sem - I

16.01.2019

37 + 23 + 29 = 94

(3 Hours)

Total Marks: 80

- N.B. :** 1) Question No.1 is compulsory.
 2) Attempt any **THREE** from the remaining questions.
 3) Figures to the right indicate full marks.

- Q1. Write a short note on the following.
- (a) Web Mining [5]
 - (b) Decision tree [5]
 - (c) OLAP operations [5]
 - (d) ADT [5]
- Q2. (a) What is distributed databases? Explain types of Distributed Database architecture. [10]
 (b) Define Data warehouse. Explain ETL process in detail. [10]
- Q3. (a) Differentiate the following [10]
1. OODBMS vs ORDBMS
 2. OLAP Vs OLTP
- (b) Apply Apriori algorithm to the following data set to find out strong association rule with Support= 50% and Confidence=70% [10]

Transaction ID	Items
101	Sharpner, Tubelights, Fan, tape
201	Sharpner, Tubelights, Fan
301	Sharpner, Screws, capacitor
401	Sharpner, Fan, Screws
501	Tubelights, Fan, Screws

- Q4. (a) Explain all phases involved in KDD process. [10]
 (b) Define and explain Bayesian and Naïve Bayesian classification. [10]
- Q5. (a) Define Data Mining. Explain Data preprocessing techniques used in Data mining process. [10]

49738

F-Y. MCA Sem I

16.

(10)

9/12

- (b) Generates the cluster using Euclidian distance for the given dataset using K means clustering. (k=2). Consider T1 and T2 as seeds/centroids for two clusters respectively.

Item	Wt	Ht
T1	1	2
T2	1	1
T3	3	4
T4	2	3
T5	2	4

Q6. (a) Define Clustering. Describe Hierarchical clustering in detail.

(10)

(b) Explain the following

(10)

1. Associative classification.
2. Text Retrieval Methods

43738