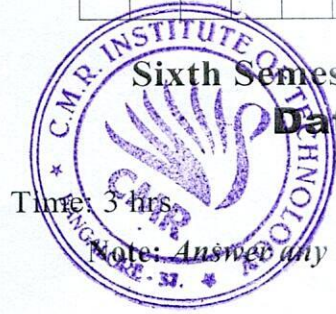


USN

18CS641



Sixth Semester B.E. Degree Examination, Dec.2024/Jan.2025

Data Mining and Data Warehousing

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- 1 a. What is data warehouse? Explain three tier architecture of data warehousing with the help of block diagram. (10 Marks)
- b. Explain different features of data warehousing. (05 Marks)
- c. Distinguish between OLTP and OLAP. (05 Marks)

OR

- 2 a. Illustrate all the conceptual modelling of data warehousing with an example for each. (10 Marks)
- b. Explain concept of hierarchy. (05 Marks)
- c. Identify different data cube measures and explain briefly. (05 Marks)

Module-2

- 3 a. Illustrate the data preprocessing steps in brief. (10 Marks)
- b. Discuss the difference between HOLAP, ROLAP, MOLAP with an example each. (10 Marks)

OR

- 4 a. Illustrate indexing of OLAP data with an example. (10 Marks)
- b. Discuss the challenges of data mining. (10 Marks)

Module-3

- 5 a. Illustrate frequent item set generation with an example by applying association rule. (10 Marks)
- b. Generate rules for the following transaction details using Apriori algorithm. Assume Minsup = 50%, MinConf = 70%.

TID	
1	{A, C, D}
2	{B, C, E}
3	{A, B, C, E}
4	{B, E}

(10 Marks)

OR

- 6 a. Illustrate Apriori algorithm for the frequent Itemset generation. (10 Marks)
- b. Explain the steps of FP – Growth algorithm. Construct FP – Tree for the below example.

FID	Item bought
100	{f, a, c, d, g, i, m, p}
200	{a, b, c, f, l, m, o}
300	{b, f, h, j, o}
400	{b, c, k, s, p}
500	{a, f, c, e, l, p, m, n}

MinSupp = 3.

(10 Marks)

Module-4

- 7 a. What is classification task? Illustrate with an example. (10 Marks)
- b. Discuss Hunt's algorithm. Construct decision tree induction using Hunt's algorithm for the following data :

TID	House owner	Marital status	Annual income	Defaulted borrower
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Divorced	95 K	Yes
5	No	Single	90 K	Yes
6	No	Married	60 K	No

(10 Marks)

OR

- 8 a. Illustrate Nearest Neighbor classifier with an example. (10 Marks)
- b. Explain sequential covering algorithm for rule based classifier. Demonstrate with an example. Assume any data set that must contain a collection of positive and negative values. (10 Marks)

CMRIT LIBRARY
BANGALORE - 560 037**Module-5**

- 9 a. Illustrate K-means algorithm with an example. (10 Marks)
- b. Explain the method for cluster evaluation and density based clustering. (10 Marks)

OR

- 10 a. Illustrate graph based clustering with an example. (10 Marks)
- b. Discuss scalable clustering algorithm with an example. (10 Marks)
