

U.S.N

--	--	--	--	--	--	--	--	--	--

P.E.S. College of Engineering, Mandya - 571 401

(An Autonomous Institution affiliated to VTU, Belgaum)

First Semester; M.Tech - Computer Engineering (MCEN)

Make-up Examination; Feb -2016

Data Warehousing and Mining

Time: 3 hrs

Max. Marks: 100

Note: Answer FIVE full questions, Selecting ONE full question from each unit.

UNIT - I

- 1 a. What is a data warehouse? Compare the features of OLTP and OLAP systems. 10
- b. Discuss the different forms of a multidimensional data model of a data warehouse? Draw a star schema of a data warehouse for books sales with 'Sales' as fact table and 'time', 'item', 'Publisher', 'branch', and 'location' as dimension tables. Assume suitable attributes for each table. 10
- 2 a. What is data mining? With a block diagram explain knowledge discovery in database. 10
- b. Discuss with example, two major categories of data mining tasks. List the challenges that monitored the development of data mining. 10

UNIT - II

- 3 a. Consider the training examples given in Table (3a). For a binary classification problem,
 - a) Compute the Gini index for the overall collection of training examples.
 - b) Compute the Gini index for the customer ID attribute. 10
 - c) Compute the Gini index for the Gender attribute.
 - d) Compute the Gini index for the car type attribute using multi way split.
 - e) Which attribute is better, Gender, Car type?
- b. Define classification with an example. Explain the general approach to solving a classification problem. 10
- 4 a. Explain when and how rule – based classifier and nearest classifier is used for classification. Illustrate with example? 10
- b. Write and explain algorithm for generating the topology of Bayesian Networks. 10

UNIT - III

- 5 a. Write and explain the Apriori algorithm for finding frequent item sets. 10
- b. Explain any two alternative methods for generating frequent item sets. 10
- 6 a. Explain the different steps involved in subsequent extension of FP – tree with example. 10
- b. Analyse, how association patterns are evaluated? 10

UNIT - IV

- 7 a. Explain the following clustering algorithm ; 10
 - i) K – means
 - ii) DBSCAN

- b. What is Agglomerative Hierarchical clustering? Explain the various techniques of Hierarchical clustering. 10
- 8 a. What is spatial data mining? Explain how spatial data cube can be constructed. 10
- b. Illustrate with an example mining of complex data objects. 10

UNIT - V

- 9 a. What is multimedia data mining? List the different multimedia data mining methods and explain any one of them. 10
- b. What is web mining? Explain the techniques used in web mining. 10
- 10 a. Consider a scenario of data mining and discuss the social impact in chart system. 10
- b. List and explain features that are considered to choose a data mining system. 10

Table 3(a) Data set

Customer ID	Gender	Car – type	Shirt size	Classes
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra large	C0
6	M	Sports	Extra large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C1
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1