Roll No. | | | | | | | | | |

108

2

COER University

END SEMESTER EXAMINATION, EVEN SEM 2022-23

Time : 3 hours
Program Name : MCA
Course Name : Data Mining and Data Warehousing

Total Marks : 100
Semester : II
Course Code : MCA213

Note: All questions are compulsory. No student is allowed to leave the examination hall before the completion of the time.

Q. No 1	Attempt Any Four Parts. Each Question Carries 5 Marks.	CO	BL
(a)	Explain the concept of Data Mining. Also explain the knowledge discovery process.	CO 1	2
(b)	Data mining as a step in the process of knowledge discovery. Justify this statement.	CO 1	2
(c)	Apply the two methods below to normalize the following group of data: 200, 300, 400, 600, and 1000. a) Use min-max normalization by setting min=0 and max=1. b) Z-score normalization	CO 1	3
(d)	Discuss the Bar chart method of data visualization.	CO 1	3
(e)	Define data transformation with example. Also define class comparison concept.	CO 1	2

Q. No 2	Attempt Any Four Parts. Each Question Carries 5 Marks.	CO	BL
(a)	Write the short notes on the Multilevel and Multidimensional Association rule mining.	CO 2	2
(b)	Discuss in detail about the Quantitative Association Rule Mining.	CO 2	2
(c)	List the various association rule mining techniques.	CO 2	1
(d)	Discuss Transactional Databases with the help of an example.	CO 2	2
(e)	Consider the data Set D. Given the minimum support 2, apply Apriori Algorithm on this dataset. **Transaction ID** — **Items** 100 — A,C,D 200 — B,C,E 300 — A,B,C,E 400 — B,E **a)** All Frequent item sets in database D using Apriori Algorithm. **b)** Strong Association rules for database D.	CO 2	3

Q. No 3	Attempt Any Four Parts. Each Question Carries 5 Marks.	CO	BL
(a)	Describe the data classification process with a neat diagram.	CO 3	4
(b)	How does the Naive Bayesian classification works? Explain with the help of example.	CO 3	2
(c)	Define Agglomerative and divisive hierarchical clustering algorithm with example.	CO 3	2
(d)	Discuss the key issue in hierarchical clustering algorithm.	CO 3	4
(e)	Define K-mean algorithm. Generate two clusters (K=2) with K-mean algorithm using data (185,72), (170,56), (168,60), (179,68), (182,72), (188,77), (180,71), (180,70), (183,84), (180,84), (180,67), (177,76) where first value is height and second value are weight.	CO 3	3

Q. No 4	Attempt Any Two Parts. Each Question Carries 10 Marks.	CO	BL
(a)	Define statistical data analysis. Also define the process of statistical method in predictive modeling along with its benefits.	CO 4	2
(b)	Explain the real-life applications of predictive modeling with proper justification.	CO 4	2
(c)	Analyze the working principle of logistic regression with mathematical modeling.	CO 4	4

108

Q. No 5	Attempt Any Two Parts. Each Question Carries 10 Marks.	CO	BL
(a)	Describe Data warehouse. Also define the architecture of data warehouse with advantages and disadvantages.	CO 5	2
(b)	Write the short note on OLAP function: a) Roll-up b) Drill-down c) Slicing d) Dicing	CO 5	2
(c)	Write in brief about schemas in multidimensional data model and discuss the applications of multidimensional data models.	CO 5	2

----------End of Paper----------