



Seventh Semester B.E. Degree Examination, Dec.2023/Jan.2024
Data Warehousing and Data Mining

Time: 3 hrs.

Max. Marks:100

Note: Answer any FIVE full questions, selecting atleast TWO questions from each part.

PART – A

- 1 a. Explain ODS as proposed by Immon and Imhoff with a possible ODS structure. (08 Marks)
b. Explain the issues resolved by ETL system. (06 Marks)
c. What is Meta data? Explain it in the context of data warehouse. (06 Marks)
- 2 a. What are the different types of OLAP operations? Explain them with suitable example. (10 Marks)
b. What is Data cube? With figure, explain different structure of data cubes. (10 Marks)
- 3 a. What is Data Mining? Explain the core data mining tasks with example. (08 Marks)
b. What is Data preprocessing? List and explain the different preprocessing techniques. (08 Marks)
c. Calculate the SMC and Jaccard similarity coefficients for the binary vector given :
 $X = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$; $Y = (0, 0, 0, 0, 0, 0, 1, 0, 0, 1)$. (04 Marks)
- 4 a. With suitable example illustrate the Apriori principle. (10 Marks)
b. Explain frequent itemset generation in FP-growth algorithm. (10 Marks)

PART – B

- 5 a. What is Classification? Explain the general approach for building classification model. (10 Marks)
b. What is Rule based Classifier? Explain how rule based classifier works. (10 Marks)
- 6 a. Estimate conditional probabilities of continuous attribute by Naïve Baye's classifier. (10 Marks)
b. Explain in detail Bagging and Boosting accuracy of classifier. (10 Marks)
- 7 a. Explain different methods of computing distance. (10 Marks)
b. Give the taxonomy of cluster analysis methods. (10 Marks)
- 8 Write short notes on :
 - i) Multiclass problem.
 - ii) Mining raster database.
 - iii) Automatic classification of web document.
 - iv) Construction of multilayered web information base. (20 Marks)

CMRIT LIBRARY
BANGALORE - 560 037

* * * * *