



USN

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

10IS74

Seventh Semester B.E. Degree Examination, Aug./Sept. 2020
Data Warehousing and Data Mining

Time: 3 hrs.

Max. Marks:100

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

PART – A

- 1 a. What is Operational Data Store (ODS)? Explain the ODS structure with neat diagram. (08 Marks)
b. What is ETL process? Explain the steps involved in data cleaning. (08 Marks)
c. What is Meta data? Explain in brief. (04 Marks)
- 2 a. What is data cube? Explain its operations with suitable examples. (10 Marks)
b. Bring out the difference between OLAP and OLTP. (05 Marks)
c. Write short note on : i) MOLAP ii) FASMI characteristics. (05 Marks)
- 3 a. Discuss the task of data mining with suitable example. (10 Marks)
b. For the vectors given below find :
i) SMC ii) Jacquard co-efficient iii) Hamming distance.
 $P = 1000000000$
 $q = 0000001001$. (08 Marks)
c. What are proximity measures? (02 Marks)
- 4 a. Explain the different factors affecting the complexity of Apriori algorithm. (08 Marks)
b. Explain FP growth algorithm for discovering frequent ItemSet. What are its limitations? (08 Marks)
c. Define : i) Confidence of a rule ii) Support of a rule. (04 Marks)

PART – B

- 5 a. Explain Hunt's algorithm and illustrate its working. (08 Marks)
b. What are nearest neighbor algorithm? (08 Marks)
c. Explain the two models of classification. (04 Marks)
- 6 a. What is Bayesian classifier? Theorem for classification with example. (10 Marks)
b. What is predictive accuracy of classification method? Explain different methods of estimating predictive accuracy and errors. (10 Marks)
- 7 a. Explain the K – means clustering algorithm with suitable example. (10 Marks)
b. Discuss about hierarchical clustering algorithm. (05 Marks)
c. What is density based clustering? (05 Marks)
- 8 Write short notes on :
a. Spatial data mining
b. Text clustering
c. Temporal data mining
d. Web content mining. (20 Marks)

* * * * *

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg. 42+8 = 50, will be treated as malpractice.