



USN

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

10IS74

**Seventh Semester B.E. Degree Examination, July/August 2021**  
**Data Warehousing and Data Mining**

Time: 3 hrs.

Max. Marks:100

**Note: Answer any FIVE full questions.**

- 1 a. Define ODS (Operational Data Store) and its structure with a neat diagram. (07 Marks)  
b. Define ETL (Extraction Transformation Loading) and functions of ETL. (07 Marks)  
c. Compare the OLTP and data warehouse system. (06 Marks)
- 2 a. Define OLAP (Online Analytical Processing) and describe CODD's OLAP characteristics. (12 Marks)  
b. Define the following data cube operations  
i) Roll-up ii) Drill down iii) Slice and Dice iv) Pivot. (08 Marks)
- 3 a. Explain the process of Knowledge Discovery in Databases (KDD). (08 Marks)  
b. Define the following :  
i) Simple matching coefficient  
ii) Jaccard similarity coefficient  
iii) Cosine similarity  
iv) Euclidean distance. (12 Marks)
- 4 a. Explain frequent itemset generation in Apriori Algorithm. (10 Marks)  
b. Define FP Growth Algorithm. Define the FP Tree representation for the following data sets :

TID	1	2	3	4	5
Items	{a, b}	{b, c, d}	{a, c, d, e}	{a, b, e}	{a, b, c}
TID	6	7	8	9	10
Items	{a, b, c, d}	{a}	{a, b, c}	{a, b, d}	{b, c, d}

 (10 Marks)
- 5 a. What is classification? Explain the two classification model with example. (10 Marks)  
b. Define Decision Tree and how a Decision Tree works. (10 Marks)
- 6 a. Explain Nearest Neighbor classifier, with algorithm. Discuss the characteristics of Nearest Neighbor classifier. (12 Marks)  
b. Define Bayesian classifier. (08 Marks)
- 7 a. What is Cluster analysis? Define the types of cluster analysis methods. (12 Marks)  
b. Define K-means Algorithm. (08 Marks)
- 8 Write short notes on :  
a. Web content mining  
b. Web usage mining  
c. Spatial data mining  
d. Data mining applications. (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.