

CBCS SCHEME

18CS641

USN

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

Sixth Semester B.E. Degree Examination, Jan./Feb. 2023 Data Mining and Warehousing

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- Explain with diagram, A Three – Tier Data warehouse Architecture. (10 Marks)
 - What is data warehouse? Explain its key features. (05 Marks)
 - Explain any five differences between OLAP and OLTP. (05 Marks)

OR

- Explain operations of OLAP. (10 Marks)
 - Explain star, snow and Fact constellation schemes. (10 Marks)

Module-2

- Explain Indexing OLAP Data : Bit Map index and join Index with example. (10 Marks)
 - Explain characteristics of Data sets. (05 Marks)
 - Explain measurement and Data collection Errors. (05 Marks)

OR

- What is Data Mining? Explain motivating challenges. (05 Marks)
 - Explain Data preprocessing Techniques. (10 Marks)
 - Find the SMC and Jaccard similarity coefficient for given two binary vectors.
 $X = \{0, 1, 0, 0, 0, 0, 0, 0, 0, 0\}$ and $Y = \{1, 0, 0, 0, 0, 0, 1, 0, 0, 0\}$. (05 Marks)

Module-3

- Explain Rule Generation in Aprior Algorithm. (10 Marks)
 - Explain Frequent item set generation of Aprior Algorithm. (10 Marks)

OR

- Explain alternative methods for generating frequent itemsets. (10 Marks)
 - A Database has five transactions. Let the minimum support be 3.
 - Find the order items set
 - Construct FP – Tree
 - Find conditional frequent pattern and frequent pattern generation by FP algorithm.

TID	ITEMS
T1	{M, O, N, K, E, Y}
T2	{D, O, N, K, E, Y}
T3	{M, A, K, E}
T4	{M, U, C, K, Y}
T5	{C, O, O, K, I, E}

(10 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.

Module-4

- 7 a. Explain the steps to build Decision Tree using Hunts Algorithm. (10 Marks)
b. Explain K-Nearest neighbor classification algorithm with example. (10 Marks)

OR

- 8 a. Explain sequential covering algorithm with an example. (10 Marks)
b. Explain Decision tree Induction algorithm with an example. (10 Marks)

Module-5

- 9 a. Explain briefly Agglomerative Hierarchical Clustering with example. (10 Marks)
b. Explain DBSCAN Algorithm with example. (10 Marks)

OR

- 10 a. Briefly explain BIRCH scalable clustering algorithm. (10 Marks)
b. Explain DENCLUE algorithm with example. (10 Marks)
