TE comp| sem Ⅴ| R-19| ATKT |FH 2023| 31|05|2023

QP code: 27279

Time: 3 hours　　　　　　　　　　　　　　　　　　　　　　　**Max. Marks: 80**

Note: 1. Question no.1 is compulsory.
2. Attempt any three out of remaining five.
3. Assumptions made should be clearly indicated.
4. Figures to the right indicates full marks.
5. Assume suitable data whenever necessary.

| Question 1 | Solve any four. | 5 marks each |

A　　What are the basic building blocks of Data warehouse?

B　　Explain Page Rank technique in detail.

C　　Compare OLTP and OLAP.

D　　Differentiate between Agglomerative and Divisive clustering method.

E　　Discuss data visualization Technique.

F　　Explain issues in Data mining.

Question 2　　　　　　　　　　　　　　　　　　　　　　　**10 marks each**

A　　Explain Decision Tree based Classification Approach with example. Discuss Metrics for evaluating Classifier Performance.

B　　Describe the steps involved in Data Mining when viewed as a process of Knowledge Discovery.

Question 3　　　　　　　　　　　　　　　　　　　　　　　**10 marks each**

A　　Differentiate between Star schema and Snowflake schema. Design Star schema for company sales with three dimensions such as Location, Item and Time.

B　　Explain Data Pre-processing.

Question 4　　　　　　　　　　　　　　　　　　　　　　　**10 marks each**

A　　Differentiate between top-down and bottom-up approaches for building data warehouse. Discuss the merits and limitations of each approach. Also explain the practical approach for designing a data warehouse.

B　　What is Web mining? Explain Web structure Mining and Web Usage Mining in detail.

Question 5   **10 marks each**

A   Explain multilevel and multidimensional association rule mining in detail.

B   A database has five transactions. Let minimum support count = 2 and minimum confidence =60 %. Find all frequent item sets using Apriori Algorithm. List strong association rules.

TID	Items
100	1,3,4
200	2,3,5
300	1,2,3,5
400	2,5
500	1,3,5

Question 6   **10 marks each**

A   Explain K-Means clustering algorithm. Discuss its advantages and limitations. Apply K-Means algorithm for the following data set with 3 clusters.

Data Set={2,3,6,8,9,12,15,18,22}

B   Consider the data given below. Create adjacency matrix. Apply complete link algorithm to cluster the given data set and draw the dendogram.

	A	B	C	D	E
A	0	2	6	10	9
B	2	0	3	9	8
C	6	3	0	7	5
D	10	9	7	0	4
E	9	8	5	4	0

_____

917952CCA45BF11A11C1029623E0917B