design the dendrogram and clusters using hierarchical clustering.

	A	B	C	D	E	F
A	**0.00**					
B	0.72	**0.00**				
C	6.66	5.95	**0.00**			
D	4.61	3.92	3.24	**0.00**		
E	5.24	4.54	2.41	2.00	**0.00**	
F	4.20	3.50	3.50	0.60	2.21	**0.00**

(CO4)

(c) Develop an algorithm for classification using decision tree. Illustrate the algorithm with a relevant example. (CO4)

5. (a) What is multimedia data mining ? How similarity search can be performed on multimedia data ? Describe the contents of a multimedia data cube. (CO5)

(b) Write short notes on the following :

(i) Multimedia data base

(ii) Data Mining Applications (CO5)

(c) What is text mining ? Also explain text data analysis and information retrieval. (CO5)

H Roll No. .....

TBC–604(3)

B. C. A. (SIXTH SEMESTER)

END SEMESTER

EXAMINATION, June, 2023

DATA WAREHOUSING AND DATA MINING

Time : Three Hours

Maximum Marks : 100

Note : (i) All questions are compulsory.

(ii) Answer any *two* sub-questions among (a), (b) and (c) in each main question.

(iii) Total marks in each main question are **twenty**.

(iv) Each sub-question carries 10 marks.

1. (a) Discuss the various characteristics of Data warehouse. Explain in three-tier architecture of data warehouse. (CO1)

P. T. O.

(b) Explain the different schemas for multidimensional databases. (CO1)

(c) Describe data mart and Meta data. Explain the Classification of Meta data. (CO1)

2. (a) What is data pre-processing ? Why Data pre-processing is so important ? What are the major tasks in Data Pre-processing ? (CO2)

(b) Explain various operation of OLAP with example. (CO2)

(c) Write short notes on the following :

(i) OLAP *vs.* OLAP

(ii) OLAP guidelines (CO2)

3. (a) What is Data Mining ? Explain the process of knowledge discovery in data base (KDD) with a neat diagram. (CO3)

(b) What is the use of data mining task ? What are the basis types of data mining tasks ? Explain with examples. (CO3)

(c) Apply the Apriori algorithm for discovering the frequent item sets of the

following table. Let the minimum support value is 2 :

T-id	Item Purchased
101	Milk, Bread, Eggs
102	Milk, Juice
103	Juice, Butter
104	Milk, Bread, Eggs
105	Coffee, Eggs
106	Coffee
107	Coffee, Juice
108	Milk, Bread, Cookies, Eggs
109	Cookies, Butter
110	Milk, Bread

(CO3)

4. (a) What is clustering ? How is this different from classification ? Discuss the applications of cluster analysis. What are the requirements of clustering in data mining ? (CO4)

(b) Write the algorithm for hierarchical clustering. For the given proximity matrix