Graphic Era HILL UNIVERSITY
Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

End Term (Even) Semester Examination May-June 2025

Roll no.....

Name of the Program and semester: BCA 6
Name of the Course: Data Warehousing and Data Mining
Course Code: TBC 604(1)

Time: 3 hours Maximum Marks: 100

Note:

(i) All the questions are compulsory.
(ii) Answer any two sub questions from a, b and c in each main question.
(iii) Total marks for each question is 20 (twenty).
(iv) Each sub-question carries 10 marks.

Q1. (2X10=20 Marks)

a. Explain the important features and the need for a Data Warehouse for businesses. **[CO1]**
b. Explain the major components of a Data Warehouse Architecture with a neat diagram. **[CO1]**
c. Compare and contrast the different schemas of the Multidimensional Model. **[CO1]**

Q2. (2X10=20 Marks)

a. Explain the different OLAP operations with suitable examples. **[CO2]**
b. Differentiate between Bitmap Indexing and Join Indexing techniques. **[CO2]**
c. Compare and contrast Multidimensional and Multirelational OLAP. **[CO2]**

Q3. (2X10=20 Marks)

a. Explain the main steps in the Knowledge Discovery from Data (KDD) process. **[CO3]**
b. Explain Association Rule mining with a suitable example. **[CO3]**
c. Consider the following dataset and find the support and confidence for the rule **{Milk, Diaper}-→{Beer}**. Minimum support is 2, and confidence is 100%. **[CO3]**

TID	Items
1	Bread, Milk
2	Bread, Milk, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Milk, Bread, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Q4. (2X10=20 Marks)

a. Explain the Decision tree classifier in detail with a suitable example. **[CO4]**
b. Classify the following examples using the K Nearest Neighbour algorithm and Euclidean distance, considering K as 2. **[CO4]**

A1= (2,10), A2= (2,5), A3= (8,4), A4= (5,8), A5= (7,5), A6= (6,4), A7= (1,2), A8= (4,9).

c. Use single and complete link agglomerative clustering to group the data described by the following distance matrix. **[CO4]**

	A	B	C	D
A	0	1	4	5
B		0	2	6
C			0	3
D				0

Graphic Era HILL UNIVERSITY
Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

End Term (Even) Semester Examination May-June 2025

Q5. (2X10=20 Marks)

a. Explain the different types of complex data with suitable examples. **[CO5]**
b. Explain in detail the challenges involved in mining Multimedia Data. **[CO5]**
c. Write short notes on a) Text Data Mining, b) Web Data Mining. **[CO5]**