Name: **Enrolment No:**	

UPES

End Semester Examination, December 2024

Course: Data Warehouse & Data Mining **Semester: VII**

Program: B.Tech (CSE) **Time: 03 hrs.**

Course Code: CSBA4022P **Max. Marks: 100**

Instructions: Please attempt according to the provided time and given weightage.

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	Why do we need normalization?	**4**	**CO1**
Q 2	Define Staging. What are the objectives of the Staging?	**4**	**CO1**
Q 3	Elaborate on the concept of Relevance and Thumb Rule for Data Understanding.	**4**	**CO2**
Q 4	What is the concept of Neighborhoods with examples?	**4**	**CO2**
Q 5	Mention are the stages involved in the evolution of data mining?	**4**	**CO1**
	SECTION B **(4Qx10M= 40 Marks)**		
Q 6	Discuss the components and key operations involved in data mining methods.	**10**	**CO1**
Q 7	Mention two techniques used for model selection. Explain with examples of different steps involved in the data cleaning process.	**2+8**	**CO2**
Q 8	Mention the key research challenges for KDD in the Knowledge Discovery in Databases (KDD) process.	**10**	**CO3**
Q 9	What do you mean by CRISP-DM? Explain the different phases involved in CRISP-DM.	**2+8**	**CO2**
	SECTION-C **(2Qx20M=40 Marks)**		
Q 10	a) For a relational scheme R (A, B, C, D, E) and set of functional dependencies F= {A→BC, CD→E, B→D, E→A}. i. Determine the closer of the given set of functional dependencies. ii. List out the candidate keys of R. iii. Identify the prime attributes and non-prime attributes. b) What are the stages involved for the evolution of data mining?	**4+8+3** **5**	**CO3** **CO2**

	Or a) What are potential issues with model selection? b) Explain the different types of classification algorithms with examples.	**8+12**	**CO3** **CO2**
Q 11	a) Discuss Star and Snowflake Schemas with comparison and examples of multi-dimensional model structures. b) Explain the different steps involved in the KDD process.	**10+10**	**CO2**