07/01/20025 MCA (NEP 2020) SEM-I ADMS QP CODE: 10071484

(2 Hours) **Total Marks: 50**

Note:
- **Question number Q1 is compulsory**
- **Attempt any two questions out of Q2 to Q5**

			Marks	Course Outcome CO	Bloom's Level BL
Q1	**Attempt any 4 from a to f**				
	a.	Object oriented Vs Object relational database.	**[05]**	CO1	BL1
	b.	What is OLAP and how does it support decision-making in data warehouses?	**[05]**	CO2	BL2
	c.	What are the metrics used to evaluate the strength of an association rule?	**[05]**	CO3	BL1
	d.	What is a parallel database system, and how does it differ from a distributed database system?	**[05]**	CO1	BL2
	e.	What is the Naive Bayes algorithm? How does the Naive Bayes algorithm work?	**[05]**	CO4	BL2
	f.	Differentiate Data Mining and Data Warehouse	**[05]**	CO4	BL2
Q2	**a.**	Define Distributed Database. Explain different types of distributed database architectures in detail.	**[08]**	CO1	BL2
	b.	Apply Agglomerative Clustering using Single Linkage for the distance matrix given below:	**[07]**	CO4	BL3,BL4 ,BL5

	A	B	C	D	E
A	0				
B	1	0			
C	4	5	0		
D	7	6	3	0	
E	9	8	5	2	0

Q3	**a.**	Use the Apriori algorithm to find frequent itemsets and association rules with a minimum support of 50% and minimum confidence of 70%.	**[08]**	CO3	BL3,BL4 ,BL5

Transaction Id	Items Bought
T1	{A, B,C}
T2	{A, B}
T3	{A, C}
T4	{B, C}
T5	{A, B,C}

	b.	What is data reduction, and why is it important in data mining and machine learning? Discuss Data cube aggregation, dimensionality reduction techniques in detail.	**[07]**	CO4	BL2

Q4	**a.**	Discuss OLAP models? How do they differ in terms of their underlying architecture?	**[08]**	CO2	BL2
	b.	What is web content mining? What is the role of web crawler in web content mining?	**[07]**	CO4	BL4
Q5	**a.**	What is the ID3 algorithm? How does the ID3 algorithm calculate entropy and information gain?	**[08]**	CO4	BL3,BL4 ,BL5

Calculate the **Entropy** of the given dataset.

Outlook	Temperature	Humidity	Wind	Play Tennis	Outlook
Sunny	Hot	High	Weak	No	Sunny
Sunny	Hot	High	Strong	No	Sunny
Overcast	Hot	High	Weak	Yes	Overcast
Rain	Mild	High	Weak	Yes	Rain
Rain	Cool	Normal	Weak	Yes	Rain
Rain	Cool	Normal	Strong	No	Rain
Overcast	Cool	Normal	Strong	Yes	Overcast
Sunny	Mild	High	Weak	No	Sunny
Sunny	Cool	Normal	Weak	Yes	Sunny
Rain	Mild	Normal	Weak	Yes	Rain

	b.	What is a Data Warehouse? How is it different from a Database? What is the significance of Dimensional Modeling in Data Warehousing?	**[07]**	CO2	BL3

****