Code No: R193204F **R19** **SET - 1**

III B. Tech II Semester Regular Examinations, June-2022

DATA MINING

(Electronics and Communication Engineering)

Time: 3 hours Max. Marks: 75

Answer any **FIVE** Questions **ONE** Question from **Each unit**
All Questions Carry Equal Marks
*****

UNIT-I

1.	a)	Explain about the major issues in data mining.	[8M]
	b)	What is the need of a data warehouse? Explain.	[7M]
		(OR)	
2.	a)	Discuss about the different types of attributes.	[8M]
	b)	Write short notes on dissimilarity between objects.	[7M]
		UNIT-II	
3.	a)	What is data Integration? What are various issues associated while performing data integration process?	[8M]
	b)	Explain data discretization techniques.	[7M]
		(OR)	
4.	a)	Explain data cube aggregation and attribute subset selection strategies for data reduction.	[8M]
	b)	Discuss data reduction as part of data preprocessing.	[7M]
		UNIT-III	
5.	a)	What is decision tree? Write and explain the algorithm for generating decision tree with an example?	[8M]
	b)	Explain Bayesian Classification.	[7M]
		(OR)	
6.	a)	What is attribute selection measure and write short notes on Gini index measure.	[8M]
	b)	Write an algorithm for decision tree induction.	[7M]
		UNIT-IV	
7.	a)	Discuss the frequent item sets generation in FP growth algorithm.	[8M]
	b)	Explain about confidence based pruning.	[7M]
		(OR)	
8.	a)	Write short notes on support counting using a hash tree with a sketch.	[8M]
	b)	Write short notes on Frequent item sets and Closed Item sets.	[7M]

UNIT-V

9. a) Discuss about the strengths and weaknesses of DBSCAN algorithm. [8M]
   b) Explain about hierarchal clustering. [7M]

(OR)

10. a) Discuss about basic K-means algorithm. [8M]
    b) Explain about different clustering techniques. [7M]

*****

Code No: R193204F **R19** **SET - 2**

III B. Tech II Semester Regular Examinations, June-2022

DATA MINING

(Electronics and Communication Engineering)

Time: 3 hours Max. Marks: 75

Answer any **FIVE** Questions **ONE** Question from **Each unit**
All Questions Carry Equal Marks
*****

UNIT-I

1.	a)	Explain about the applications of data mining.	[8M]
	b)	Discuss the data visualization.	[7M]
		(OR)	
2.	a)	Write the need and usage of data mining technologies.	[8M]
	b)	Explain about the similarity between simple attributes.	[7M]
		UNIT-II	
3.	a)	Explain in detail about various data transformation techniques used in data preprocessing.	[8M]
	b)	Describe how the data can be integrated.	[7M]
		(OR)	
4.	a)	Briefly discuss about data cube aggregation with an example in data reduction process.	[8M]
	b)	What is the need of data preprocessing and explain about different preprocessing techniques.	[7M]
		UNIT-III	
5.	a)	What is Decision Tree Induction? Explain how it classifies the data.	[8M]
	b)	Write generic Decision Tree Induction Algorithm.	[7M]
		(OR)	
6.	a)	Explain Attribute selection measures using Information gain with a suitable example.	[8M]
	b)	Write about predicting a class label using Naive Bayesian classification.	[7M]
		UNIT-IV	
7.	a)	Explain about different types of candidate generation procedures.	[8M]
	b)	Write an algorithm for rule generation using apriori.	[7M]
		(OR)	
8.	a)	Discuss about maximal frequent item sets with an example.	[8M]
	b)	Explain about FP-Growth algorithm.	[7M]

UNIT-V

9. a) Discuss different types of clusters. [8M]
   b) Explain about the issues related to k-means algorithm. [7M]

(OR)

10. a) Write short notes on agglomerative hierarchal clustering. [8M]
    b) Explain about the strengths and weaknesses of k-means algorithm. [7M]

*****

Code No: R193204F **R19** **SET - 3**

III B. Tech II Semester Regular Examinations, June-2022

DATA MINING

(Electronics and Communication Engineering)

Time: 3 hours Max. Marks: 75

Answer any **FIVE** Questions **ONE** Question from **Each unit**
All Questions Carry Equal Marks
*****

UNIT-I

1. a) Explain about the different types of data and patterns to be mined. [8M]
   b) Write short notes on pattern recognition. [7M]

(OR)

2. a) Differentiate between Database Query and Mining. [8M]
   b) Discuss about different types of attributes. [7M]

UNIT-II

3. a) What is a noise? What are the various data smoothing techniques? [8M]
   b) Describe various methods for handling tuples with missing values for some attributes. [7M]

(OR)

4. a) Write about data integration and transformation. [8M]
   b) Write about Numerosity Reduction. [7M]

UNIT-III

5. a) What is decision tree? Explain how does it work for classification problem? [8M]
   b) Discuss the Naïve Bayesian Classification with an example. [7M]

(OR)

6. a) Write an algorithm for decision tree induction. [8M]
   b) Discuss about the following (i) Gini index (ii) Information Gain [7M]

UNIT-IV

7. a) Discuss about candidate generation and pruning with an example. [8M]
   b) Explain about closed frequent item sets with an example. [7M]

(OR)

8. a) Discuss about the construction of a FP tree with an example. [8M]
   b) Explain about the support and confidence with examples. [7M]

Code No: R193204F **R19** **SET - 3**

UNIT-V

9.	a)	Discuss about DBSCAN algorithm.	[8M]
	b)	Write short notes on bisecting k-means algorithm.	[7M]
		(OR)	
10.	a)	Write short notes on different clustering techniques.	[8M]
	b)	List out the strengths and weaknesses of hierarchal clustering.	[7M]

*****

Code No: R193204F **R19** **SET - 4**

III B. Tech II Semester Regular Examinations, June-2022

DATA MINING

(Electronics and Communication Engineering)

Time: 3 hours Max. Marks: 75

Answer any **FIVE** Questions **ONE** Question from **Each unit**

All Questions Carry Equal Marks

*****

UNIT-I

1.	a)	What is the need of a data warehouse and differentiate between data warehouse and data base.	[8M]
	b)	Discuss about various visualization techniques.	[7M]
		(OR)	
2.	a)	Discuss the major issues in data mining.	[8M]
	b)	Explain about the importance of data mining technologies.	[7M]
		UNIT-II	
3.	a)	What is Data Integration? What are various issues associated while performing data integration process?	[8M]
	b)	Explain data discretization techniques.	[7M]
		(OR)	
4.	a)	What is data preprocessing? Explain why it is essential before mining the knowledge from data.	[8M]
	b)	Illustrate data reduction strategies.	[7M]
		UNIT-III	
5.	a)	Differentiate Supervised and Unsupervised Learning.	[8M]
	b)	Explain Tree Pruning in detail.	[7M]
		(OR)	
6.	a)	What is classification? How does classification work? Discuss its process steps.	[8M]
	b)	Explain about Bayes' Theorem.	[7M]
		UNIT-IV	
7.	a)	Explain about Market Basket analysis.	[8M]
	b)	Discuss about Association Rule Mining.	[7M]
		(OR)	
8.	a)	Explain about the apriori principle.	[8M]
	b)	Write an apriori algorithm for frequent item set generation.	[7M]
		UNIT-V	
9.	a)	Write short notes on k-means as an optimization problem.	[8M]
	b)	Explain about the importance of cluster analysis.	[7M]
		(OR)	
10.	a)	Discuss the key issues in hierarchal clustering.	[8M]
	b)	Explain about basic k-means algorithm.	[7M]

*****

|'''|'|'|''||'''||||