



**NARASIMHAREDDYENGINEERINGCOLLEGE**

(Autonomous)

Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad

Accredited by NAAC with A Grade, Accredited by NBA

**IV-I sem COMPUTER SCIENCE AND ENGINEERING**

**QUESTION BANK**

Course Title: Data Mining

**UNIT-I**

S.No	Questions						
<b>Part –A(Short Answer Questions)</b>							
1	What is Data Mining?						
2	What is KDD process?						
3	Explain Data cleaning method.						
4	What is Data Preprocessing?						
5	What is data reduction?						
6	Draw a neat diagram of Data Mining architecture.						
7	List different types of Data Mining tasks.						
8	Difference between Data cleaning and Data reduction.						
9	Explain data transformation method.						
10	What is Data transformation? Explain.						
<b>Part– B(Long Answer Questions)</b>							
11	a)	What is KDD process? Explain with neat diagram.					
	b)	Explain the Noisy and missing data.					
12	a)	What is data mining? Explain architecture of data mining.					
	b)	Explain Dimensionality reduction process with diagram.					
13	Explain the following: a. Data cleaning						
	b) Data transformation						
14	a)	Explain different types of data preprocessing techniques.					
	b)	Explain in detail about measures of similarity and Dissimilarity.					
15	Discuss about the following: a. missing data						
	b) Dimensionality Reduction						

**UNIT-II**

S.No	Questions					
<b>Part –A(Short Answer Questions)</b>						
1	How the Association rule is helpful to growth of business.					
2	What are disadvantages of Apriori Algorithm?					
3	What is market basket analysis? Explain.					
4	Discuss the applications of association analysis.					
5	What are the advantages of FP-Growth algorithm?					
6	Write a step by step process of FP Growth Algorithm.					
7	Write about basic concept in Association Rule Mining.					
8	Explain the process of Apriori Algorithm.					
9	What are the advantages and disadvantages of ECLAT algorithm?					

10		Explain the process of ECLAT algorithm.			
<b>Part– B(Long Answer Questions)</b>					
11	a)	Explain about the Apriori algorithm for finding frequent item sets with an example.			
	b)	Write a step by step process of Apriori algorithm.			
12	a)	Discuss about basic concepts of frequent item set mining.			
	b)	Write the advantages and disadvantages of Apriori Algorithm.			
13	a)	Write about basic concept in Association Rule Mining.			
	b)	Can we overcome the drawbacks of Apriori algorithm? Discuss.			
14		Explain the following: a. Frequent item set.			
	b)	Closed Frequent item set			
15	a)	What are the drawbacks of Apriori Algorithm? Explain.			
	b)	Write a FP Growth Algorithm and explain.			

### UNIT-III

<b>S.No</b>	<b>Questions</b>				
<b>Part –A(Short Answer Questions)</b>					
1		What is the difference between classification and prediction?			
2		What is Bayes theorem? Explain.			
3		Write a note on attribute selection measures.			
4		What is prediction?			
5		Discuss about Naïve Bayesian Classification.			
6		Draw the structure of Decision tree with example.			
7		Explain Bayesian belief network.			
8		Discuss about Accuracy and Error measures.			
9		What is classification? Explain.			
10		What are the uses of classification technique?			
<b>Part– B(Long Answer Questions)</b>					
11		Describe the data classification process with a neat diagram. How does the Naive Bayesian Classification work? Explain.			
12	a)	Describe the data classification process with a neat diagram.			
	b)	How does the Naive Bayesian Classification work? Explain.			
13		Write and explain about Classification by Backpropagation Algorithm.			
14	a)	What is Bayesian belief network? Explain in detail.			
	b)	Write a note on attribute selection measures.			
15	a)	Explain decision tree induction algorithm for classifying data tuples and discuss suitable Example.			
	b)	Explain construction of decision tree with example.			

### UNIT-IV

<b>S.No</b>	<b>Questions</b>				
<b>Part –A(Short Answer Questions)</b>					
1		Explain about Types of Data in Cluster Analysis?			
2		Differentiate between AGNES and DIANA algorithms.			
3		Write a K-means clustering algorithm.			
4		Explain Grid based clustering methods.			

5		Writethekeyissueinhierarchicalclusteringalgorithm.			
6		Whatistheuseofclustering?			
7		Explaininterclustering.			
8		Whatisclustering?Howitisusefultobusiness.			
9		Explainunsuperviseddata.			
10		Whatisintraclustering?			
<b>Part– B(LongAnswerQuestions)</b>					
11	a)	ClassifyvariousClusteringmethods.			
	b)	WriteanyonePartitioningbasedclusteringmethods.			
12	a)	Whatisthegoalofclustering?Explan.			
	b)	Howdoespartitioningaroundmedoidsalgorithmachievethis?			
13		Explainthefollowing:a)Densitybasedclusteringmethods			
	b)	Gridbasedclusteringmethods.			
14	a)	GiveabriefnoteonPAMAlgorithm.			
	b)	Whatisdrawbackof k-meansalgorithm?Howcanwemodifythe algorithm to diminish?			
15	a)	Whatisoutlierdetection?Explaindistancebasedoutlierdetection.			
	b)	Writepartitioningaroundmedoidsalgorithm.			

## Unit-V

S.No	Questions				
<b>Part –A(ShortAnswerQuestions)</b>					
1		Whatisweb mining?			
2		WhatisTextmining?			
3		Explaintextclustering.			
4		Whatiswebcontentmining?			
5		Describehierarchyofcategories			
6		Differencebetweenwebminingandtextmining.			
7		Explainthedifferentdataminingmethods.			
8		WhatisheuseofTextmining?			
9		Explaintypes oftextmining.			
10		Explainapplicationsofwebmining.			
<b>Part– B(LongAnswerQuestions)</b>					
11	a)	DiscussabouttheWebmining			
	b)	Write a note about the web content mining.			
12	a)	Explain indetail about text mining.			
	b)	WhatisTextclustering?Explainindetails.			
13	a)	Whatismining?Explainwebstructuremining.			
	b)	Explain howdataminingisusefultoDataScience.			
14		Explainthefollowing:a)Unstructuredtext			
	b)	Hierarchicalcategories.			
15		Discussabouttheepisoderulediscoveryfortexts.			

## UNIT-V