

**PRAGATI ENGINEERING COLLEGE: SURAMPALEM**  
**(AUTONOMOUS)**  
**II B.Tech II Semester Regular/Supplementary Examinations, May-2024**

**DATA MINING**  
**(Common to CSE(AI&ML) CSE(AI) and CSE(DS))**

Time: 3 hours

Max. Marks: 70

**Answer ONE Question from each Unit**  
**All Questions Carry Equal Marks**

Q. No.	Questions	BTL	CO	Marks
<b>UNIT – I</b>				
1.	a) Compare and contrast Online Transaction Processing, Online Analytical Processing.	K4	CO1	7M
	b) Briefly discuss about the following terms i)Enterprise warehouse ii)Data mart iii)Virtual warehouse	K2	CO1	7M
<b>OR</b>				
2.	a) What is Multi Dimensional Data Model?. Explain its operations with neat Diagram.	K2	CO1	7M
	b) Explain about snowflake schema, star schema in detail.	K2	CO1	7M
<b>UNIT – II</b>				
3.	a) Briefly discuss about different Data Mining Tasks.	K2	CO2	7M
	b) Explain about different measures of similarity and dissimilarity.	K2	CO2	7M
<b>OR</b>				
4.	a) Explain about Discretization and Binarization with suitable example.	K3	CO2	7M
	b) Discuss about feature subset selection and feature creation in detail.	K2	CO2	7M
<b>UNIT – III</b>				
5.	a) What is classification? Explain general approach to solve a classification problem.	K2	CO3	7M
	b) Explain about how to build a decision tree in detail.	K2	CO3	7M
<b>OR</b>				
6.	a) What is model overfitting? Explain the reasons behind the model overfitting in detail.	K2	CO3	7M
	b) Explain about hold out method and cross validation methods to evaluate the performance of a classifier.	K2	CO3	7M
<b>UNIT – IV</b>				
7.	a) What is market basket analysis?. Explain the use of support and confidence in market basket analysis.	K2	CO4	7M
	b) State apriori principle. Explain apriori algorithm with suitable example.	K3	CO4	7M
<b>OR</b>				
8.	a) What is FP tree? Explain how to generate frequent itemsets using FP growth algorithm.	K2	CO4	7M
	b) Write a short note on maximal frequent itemset and closed frequent itemset.	K2	CO4	7M
<b>UNIT – V</b>				
9.	a) What is clustering analysis? Discuss about different types of clustering in detail.	K2	CO5	7M
	b) Discuss about strengths and weaknesses of K means clustering algorithm in detail.	K2	CO5	7M
<b>OR</b>				
10.	a) Explain about agglomerative hierarchical clustering in detail.	K2	CO5	7M
	b) Explain about DBSCAN algorithm with its strengths and weaknesses.	K2	CO5	7M