Course Code: 20MC9118 **R20**

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)
Siddharth Nagar, Narayanavanam Road – 517583
QUESTION BANK (DESCRIPTIVE)

Subject with Code: Data Warehousing and Data Mining (20MC9118) **Course & Branch**: MCA

Regulation: R20 **Year & Sem:** II-MCA & I-Sem

UNIT–I

Introduction, Data Preprocessing

1	a)	What are the major issues in data mining?	[L1][CO1]	[06M]
	b)	Discuss about the classification of data mining systems.	[L2][CO1]	[06M]
2		Generalize data preprocessing techniques.	[L2][CO1]	[12M]
3		Explain the following concepts i) Relational databases ii) Outlier analysis iii) Numerosity reduction	[L2][CO1]	[12M]
4		How to Mine frequent patterns and associations? Explain.	[L2][CO2]	[12M]
5	a)	What is data mining? Explain data mining task primitives?	[L2] [CO1]	[06M]
	b)	Identify and Demonstrate data reduction techniques.	[L3][L2][CO1]	[06M]
6	a)	Describe data integration and transformation?	[L2][CO1]	[06M]
	b)	Examine the data cleaning as a process and correlations.	[L3][CO1]	[06M]
7		Explain advanced data and information systems with advanced applications.	[L2][CO1][CO2]	[12M]
8		Analyze the following concepts in detail. i) Noisy data ii) Attribute sub set selection iii) Cluster analysis iv) Classification and prediction	[L4][CO1][CO3]	[12M]
9	a)	Discuss in detail Data Mining functionalities.	[L2][CO1]	[06M]
	b)	Explain data discretization and concept hierarchy generation.	[L2][CO1]	[06M]
10	a)	Describe data warehouses and its importance.	[L2][CO1]	[06M]
	b)	Illustrate the classification of Data Mining Systems.	[L3][CO1][CO2]	[06M]

Course Code: 20MC9118 **R20**

UNIT–II

Data Warehouse and OLAP Technology for Data Mining, Data Cube Computation and Data Generalization

1	a)	Differentiate between Operational Database Systems and Data Warehouses.	[L2][CO1]	[06M]
	b)	How data can be modeled and viewed in multiple Dimensions? Explain with Example.	[L2][CO3][CO5]	[06M]
2		Examine different schemas for multi-dimensional Data Model.	[L3][CO3][CO5]	[12M]
3		Design and elaborate the dataware house architecture.	[L3][CO1]	[12M]
4	a)	Explain concept hierarchy with an example.	[L2][CO1]	[06M]
	b)	Sketch the architecture of OLAM.	[L3][CO1]	[06M]
5		Illustrate OLAP operations with examples for each..	[L3][CO1]	[12M]
6		Summarize the data cube computation and data generalization?	[L2][CO3]	[12M]
7	a)	Examine the constrained gradient analysis in data cubes.	[L3][CO3]	[06M]
	b)	How to index OLAP data.	[L2][CO1][CO2]	[06M]
8		Discuss efficient methods for data cube computations.	[L2][CO3]	[12M]
9	a)	Explain mining class comparisons and class description.	[L2][CO6]	[06M]
	a)	Describe complex aggregation at multiple granularities.	[L2][CO1]	[06M]
10		What is a Data Cube? How data cubes can be used for multidimensional representation of data.	[L1][CO3]	[12M]

Course Code: 20MC9118 **R20**

UNIT–III

Mining Frequent Patterns, Associations and Correlations, Classification and Prediction

1		Explain Apriori algorithm with suitable example.	[L2][CO2]	[12M]
2	a)	Demonstrate frequent-pattern growth (FP-Growth) method.	[L2][CO2]	[06M]
	b)	Explain the generation of association rules from frequent item sets.	[L2][CO2]	[06M]
3		Analyze concepts of mining frequent patterns using Market Basket Analysis.	[L4][CO2]	[12M]
4	a)	Explain classification by decision tree induction.	[L2][CO2]	[06M]
	b)	Explain classification by back propagation.	[L2][CO2]	[06M]
5		Describe the Bayesian and rule based classification with examples.	[L2][CO2][CO3]	[12M]
6	a)	Analyze the issues regarding classification and prediction.	[L4][CO2]	[06M]
	b)	Explain the need and importance of Tree Pruning.	[L2][CO2]	[06M]
7		Examine different attribute selection measures with examples.	[L3][CO3]	[12M]
8	a)	Describe support vector machines?	[L2][CO2]	[06M]
	b)	Explain Bayesian Belief Networks.	[L2][CO2]	[06M]
9	a)	How to evaluate the accuracy of a classifier and predictor?	[L2][CO2]	[06M]
	b)	Interpret the relation between from association mining to correlation analysis.	[L3][CO6]	[06M]
10	a)	Discuss k-Nearest-Neighbor Classifiers Rough set approach	[L2][CO2]	[06M]
	b)	Differentiate between Linear Regression and Nonlinear Regression	[L4][CO2]	[06M]

Course Code: 20MC9118 **R20**

UNIT –IV

Cluster Analysis Introduction, Mining Streams

1	a)	Discuss in detail K-Means algorithm.	[L2][CO3]	[06M]
	b)	Explain the types of data in cluster analysis.	[L2][CO3]	[06M]
2		Explain the following Hierarchical Clustering methods. i) Agglomerative ii) Divisive	[L2][CO3]	[12M]
3		Illustrate different density based clustering methods.	[L3][CO3]	[12M]
4		Demonstrate CLIQUE (CLustering In QUEst) algorithm to cluster High Dimensional data.	[L2][CO3]	[12M]
5		What is an outlier? Discuss different ways to detect outlier.	[L1][CO3]	[12M]
6	a)	List and explain the methodologies for processing streaming data.	[L1][CO4]	[06M]
	b)	Explain in detail the following terms of Time Series analysis i) Trend Analysis ii) Similarity Search	[L2][CO5][CO6]	[06M]
7		How to mine sequence patterns in Transactional Databases?	[L2][CO5][CO6]	[12M]
8	a)	Deduce biological sequences and hidden Markov model?	[L4][CO5][CO6]	[06M]
	b)	Explain multi relational data mining.	[L2][CO5][CO6]	[06M]
9		Generalize the graph mining and social network analysis.	[L2][CO5]	[12M]
10	a)	Discover mining sequence patterns in biological data.	[L3][CO5]	[06M]
	b)	Identify the areas of mining social network.	[L3][CO5]	[06M]

Course Code: 20MC9118 **R20**

UNIT– V

Mining Object, Spatial, Multimedia, Text and Web Data, Applications and Trends in Data Mining

1		Discuss in detail the construction of Spatial Data Cube and Spatial OLAP.	[L2][CO5]	[12M]
2	a)	List and explain the approaches used to find similarities in Multimedia data.	[L2][CO4]	[06M]
	a)	What kind of associations can be mined in Multimedia data? Give Examples.	[L1][CO4]	[06M]
3	a)	Discuss in detail basic measures of Text Retrieval.	[L2][CO4]	[06M]
	b)	Explain different categories of Text retrieval methods.	[L2][CO4]	[06M]
4	a)	Illustrate different approaches to Text Mining.	[L3][CO4]	[06M]
	b)	Analyze various issues related to text mining.	[L4][CO4]	[06M]
5		List and explain the applications of Data Mining.	[L1][CO1]	[12M]
6	a)	What is Statistical Data Mining? Explain.	[L2][CO6]	[06M]
	b)	Represent the dimensionality reduction for text.	[L2][CO4]	[06M]
7	a)	Discuss the concept of multidimensional analysis.	[L2][CO5]	[06M]
	b)	Describe about descriptive mining of complex data objects.	[L2][CO6]	[06M]
8		Analyze the social implications of data mining.	[L4][CO1]	[12M]
9	a)	Explain briefly Visual and Audio Data Mining.	[L2][CO5]	[06M]
	b)	Illustrate collaborative filtering approach of data mining	[L3][CO5]	[06M]
10	a)	Generalize the trends in Data Mining.	[L2][CO1]	[06M]
	b)	What are the principles adopted for preserving privacy and data security	[L1][CO1]	[06M]

Prepared By:
Mr. V. Harsha Vardhan,
Associate Professor,
Dept. of MCA,
SIETK, Puttur.