SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY Siddharth Nagar, Narayanavanam Road, Puttur – 517583

QUESTION BANK (DESCRIPTIVE)

Subject with Code : Data warehousing and Mining(18MC9116) **Course & Branch**: MCA

Year & Sem: II-MCA & I-Sem **Regulation:** R18

UNIT –I

Introduction

1.	Explain Major issues in Data Mining and Classification of Data Mining	12M
2.	Explain Data Preprocessing Techniques	12M
3.	Explain the following concepts	
	(a) Relational Databases	04M
	(b) Outlier Analysis	04M
	(c) Numerosity Reduction	04M
4.	Explain Mining Frequent Patterns and Associations.	12M
5.	Explain Advanced Data and Information Systems and Advanced Applications	12M
6.	(a) What are Data Mining and Explain Data Mining Task Primitives?	06M
	(b) Explain Data Reduction Techniques.	06M
7.	(a) Describe Data Integration and Transformation?	06M
	(b) Explain Data cleaning as a process and Correlations.	06M
8.	Explain	
	(a) Noisy Data	03M
	(b) Attribute Subset Selection	03M
	(c) Cluster Analysis	03M
	(d) Classification and Prediction	03M
9.	(a) Explain Data Mining Functionalities and what Kind of pattern can be mined.	06M
	(b) Explain Data Discretization and Concept hierarchy generation.	06M
10.	(a) Describe Data Warehouses and its Importance.	06M
	(b) Explain Data Descriptive Data Summarization.	06M

UNIT –II

Data Warehouse and OLAP Technology for Data Mining

1.	(a) What is indexing OLAP data?	06M
	(b) Explain Multidimensional Data Model.	06M
2.	(a) Explain Data Warehouse Implementation.	06M
	(b) Describe Meta Data Repository?	06M
3.	Explain Data Warehouse Architecture.	12M
4.	(a) What is Star net Query Model for Querying?	06M
	(b) Describe from Data Warehousing to Data Mining	06M
5.	(a) Describe efficient methods for Data cube Computation	06M
	(b) Explain BUC: Computing Iceberg Cubes from the Apex Cuboid Downward.	06M
6.	(a) Explain Constrained Gradient Analysis in Data Cubes	06M
	(b) Write Pre computing Shell Fragments for Fast High-Dimensional OLAP.	06M
7.	(a) Describe Attribute Oriented Induction.	06M
	(b) Explain Data Generalization and Concept Description.	06M
8.	(a) Explain Mining Class Comparisons and Class Description.	06M
	(b) Describe Complex Aggregation at Multiple Granularities.	06M
9.	Explain Star-Cubing: Computing Iceberg Cubes.	12M
10.	Write Data cube computation and Data Generalization?	12M

UNITI-III

Mining Frequent Patterns, Associations and Correlations

1.	Explain basic concepts of Mining frequent patterns.	12M
2.	Explain scalable frequent item set Mining methods.	12M
3.	(a) What is Clustering methods and high dimensional data?	06M
	(b) Explain Time series and sequenced data.	06M
4.	Explain Bayesian and Rule based Classification with examples.	12M
5.	(a) Explain Classification by Decision Tree Induction.	06M
	(b) Explain Classification by Back propagation.	06M
6.	(a) Describe issues regarding classification and prediction.	04M
	(b) Explain Accuracy and Error measures.	08M
7.	Explain Ensemble Methods.	12M
8.	(a) Describe support vector Machines?	06M
	(b) Describe Associative Classification.	06M
9.	Describe Mining Data Streams and Lazy Learners.	12M
10.	(a) Explain the evaluating the accuracy of a Classifier and Predictor.	06M
	(b) Explain from Association Mining to Correlation Analysis.	06M
11.	Explain Classification and Prediction.	12M
12.	Explain	12M
	(a) Rough Set Approach	
	(b) Bayesian Belief Networks	
	(c) k-Nearest-Neighbor Classifier	
13.	Explain	12M
	(a) Genetic Algorithm	
	(b) Bootstrap	
	(c) Nonlinear Regression	

UNIT-IV

Cluster Analysis Introduction

1.	(a) What is Partitioning Methods?	04M
	(b) Explain the types of Data in Cluster Analysis?	08M
2.	(a) Explain Mining Data Streams.	06M
	(b) Describe Constraint based Cluster Analysis.	06M
3.	Explain Clustering High Dimensional Data.	12M
4.	Explain the following	
	(a) DENCLUE	04M
	(b) Wave Cluster	04M
	(c) DBSCAN	04M
5.	Explain the Mining Time-Series Data.	12M
6.	(a) What is Outlier Analysis? Explain it clearly.	06M
	(b) Describe Density based Outlier detection	06M
7.	(a) What is biological sequences and hidden markov model?	06M
	(b) Explain Multi relational Data Mining.	06M
8.	Explain the Graph Mining and Social Network Analysis	12M
9	(a) Explain Mining Sequence Patterns in Biological Data.	06M
	(b) Explain Hierarchical Methods.	06M
10.	Explain the Grid based methods.	12M

UNIT-V

Mining Object, Spatial, Multimedia, Text and Web Data

1.	(a) What is Generalization of Structural Data?	06M
	(b) Explain the Spatial Data Mining.	06M
2.	(a) Explain the Multimedia Data Mining.	06M
	(b) Describe Data Mining Applications.	06M
3.	Explain the Text Mining and its Importance.	06M
4.	(a) Explain Mining the World Wide Web.	06M
	(b) Explain Additional Themes on Data Mining.	06M
5.	Describe Social Impacts on Data Mining.	12M
6.	Explain Data Mining and Collaborative Filtering.	06M
7.	(a) Describe Generalization of Class Composition Hierarchies.	06M
	(b) Explain the Dimensionality Reduction for text.	06M

Prepared by: DR. A. SWARUPA RANI, Assoc. Professor, Dept. of MCA