SCHEME OF COURSE WORK
Department of CSE

Course Title	*: Data Warehousing and Data mining*		
Course Code	**:** 13IT2114	**L P C**	**4 0 3**
Programme:	**: M.Tech.**		
Specialization:	**: CSE**		
Semester	**:Ist Semester**		
Prerequisites	**: DBMS**		
Courses to which it is a prerequisite	**: Text Mining**		

Course Outcomes (COs):

1	Use the appropriate abstract data type for formulating solutions for the given problem.
2	Describe priority queues using heaps and compare the complexities of various sorting algorithms
3	Examine the solution for dynamic equivalence problem using find and smart union algorithms and discover solutions for various graph problems.
4	Extrapolate various algorithm design techniques with examples and compute amortized analysis for skew heaps, binomial queues, splay trees.
5	Apply various advanced data structures like red-black trees, heap, AA trees, k-d trees etc in relevant application needed.

Program Outcomes (POs):

A graduate of M.Tech CSE Specialization will be able to

1	Graduates will demonstrate knowledge in core subjects of Computer Science and Engineering and the ability to learn independently.
2	Graduates will demonstrate the ability to design a software application or process that meets desired Specifications within the constraints.
3	Graduates will demonstrate the ability to solve problems relevant to industries and research organizations.
4	Graduates will develop innovative thinking capabilities to promote research in core and trans-disciplinaryareas.
5	Graduates will be familiar with modern engineering software tools and equipment to analyze computer science and engineering problems.
6	Graduates will demonstrate the ability to collaborate with engineers of other disciplines and work on projects requiring multidisciplinary skills.
7	Graduates will acquire project management and finance control abilities.
8	Graduates will be able to communicate effectively in both verbal and written forms.
9	Graduates will engage themselves in lifelong learning in the context of rapid technological changes in computer science and engineering

10	Graduates will demonstrate an appreciation of ethical and social responsibilities in professional and societal context.
11	Graduates will demonstrate the ability in carrying out tasks independently and by reflective learning.

Course Outcome versus **Program Outcomes:**

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1		M	M	S	S	S				M	S
CO2		S	S	S	S	M	M				S
CO3		S	S	S	S	M	M				S
CO4		S	S	S	S	M					S
CO5		S	S	S	S	M					S

Assessment Methods:	Assignment / Quiz / Seminar / Case Study / Mid-Test / End Exam

Teaching-Learning and Evaluation

Week	TOPIC / CONTENTS	Course Outcomes	Sample questions	TEACHING-LEARNING STRATEGY	Assessment Method & Schedule
1	Introduction:Data mining-On what kinds of Data, Data Mining Functionalities, Classification of Data Mining systems, Data Mining Task Primitives	CO-1	1.Define Data mining. 2.Compare and contrast Data, information and knowledge.	▫ Lecture / Discussion	Assignment (Week 7-8) Mid-Test 1 (Week 9)
2	Integration of a Data Mining System with a Database or Data Warehouse System, Major issues in Data Mining.	CO-1	. 1. What is the difference between prediction and classification	▫ Lecture / Discussion	Assignment (Week 7-8) Mid-Test 1 (Week 9)
3	Data Preprocessing: Descriptive data summarization, Data Cleaning, Data Integration and Transformation,	CO-1	1.What is the need of Preprocessing.	▫ Lecture / Discussion ▫ Problem solving	Mid-Test 1 (Week 9)
4	Data Reduction, Discretization and Concept Hierarchy Generation.	CO-1	1. What are the phases in preprocessing.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 7-8) Mid-Test 1 (Week 9)
5	Data Warehouse and OLAP Technology: Multidimensional Data Model, Data WarehouseArchitecture	CO-2	1.What is the model used for construction of a warehouse. 2. Whar are the ways in which the warehouse may be coupled with the data mining system 3. Expand OLAP system 4. Give the difference between OLAP and OLTP.	▫ Lecture ▫ Problem solving	Mid-Test 1 (Week 9)

6	Data Warehouse Implementation, From Data Warehousing to Data Mining	CO-2	1.What are the application areas of OLAP system.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 7-8) Mid-Test 1 (Week 9)
7	Data Cube Computation and Data Generalization: Efficient methods for Data Cube Computation, Further Development of Data Cube and OLAP Technology, Attribute-Oriented Induction.	CO-2	1 Explain BUC algorithm.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 7-8) Mid-Test 1 (Week 9)
8	Mining Frequent Patterns, Association and Correlations: Basic concepts, Efficient and Scalable Frequent Itemset Mining Methods	CO-3	1.Differnce between association and correlation. 2. When is an item said to be frequent. 3. Define support and confidence.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 7-8) Mid-Test 1 (Week 9)
9	**Mid-Test 1**	CO-3			
10	Mining Various kinds of Association Rules	CO-3	1. What is the purpose of mining frequent item sets. 2.What re the drawbacks of apriori algorithm.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
11	From Association Mining to Correlation Analysis, Constraint Based Association	CO-3	1 What are constraints imposed over assoc rules.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
12	Classification and Prediction-1: Issues Regarding Classification and Prediction, Classification by Decision Tree Induction	CO-4	1.Give the formula for gainratio. 2. What is bayes rule. 3. Give the formula for error in back propagation classification.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
13	Bayesian Classification,Rule-Based Classification, Classification by	CO-4	1. What is the basic ides in ID3 algorithm.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)

	Backpropagation.		2.What is training set.		
14	Classification and Prediction-2: Support Vector Machines, Association Classification, Other Classification Methods	CO-4	1 What are the different types of SVM's.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
15	Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor.	CO-4	1 Give the formula for accuracy.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
16	Cluster Analysis Introduction :Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical Methods,	CO-5	1.Define cluster. 2. Give the formula for precision and recall. 3. What is clusterability.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
17	Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Outlier Analysis	CO-5	1. Mention different types of clustering techniques 1.Give example for Partional clustering.	▫ Lecture / Discussion ▫ Problem solving	Assignment (Week 15-17) Mid-Test 2 (Week 18)
18	**Mid-Test 2**				
19/20	**END EXAM**				