

**ST. ANN'S COLLEGE OF ENGINEERING & TECHNOLOGY: CHIRALA
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

LESSON PLAN

Subject: Data Ware housing and Mining
Name: Dr.A.VERASWAMY

Academic Year: 2017-18
Year & Sem/Section: III-II-SEM 'A&B'

Unit No.	Sub Topic Names	No. of classes required
1	Introduction: What Motivated Data Mining? Why Is It Important, Data Mining—On What Kind of Data, Data Mining Functionalities—What Kinds of Patterns Can Be Mined? Are All of the Patterns Interesting? Classification of Data Mining Systems, Data Mining Task Primitives, Integration of a Data Mining System with a Database or Data Warehouse System, Major Issues in Data Mining.	9
2	Data Pre-processing: Why Pre-process the Data? Descriptive Data Summarization, Data Cleaning, Data Integration and transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation.	9
3	Data Warehouse and OLAP Technology: An Overview: What Is a Data Warehouse? A Multidimensional Data Model, Data Warehouse Architecture, Data Warehouse Implementation, From Data Warehousing to Data Mining.	11
4	Classification: Basic Concepts, General Approach to solving a classification problem, Decision Tree Induction: Working of Decision Tree, building a decision tree, methods for expressing an attribute test conditions, measures for selecting the best split, Algorithm for decision tree induction. Model Over fitting: Due to presence of noise, due to lack of representation samples, evaluating the performance of classifier: holdout method, random sub sampling, and cross-validation, bootstrap.	8
5	Association Analysis: Basic Concepts and Algorithms: Introduction, Frequent Item Set generation, Rule generation, compact representation of frequent item sets, FP-Growth Algorithm.	5
6	Cluster Analysis: Basic Concepts and Algorithms: What Is Cluster Analysis? Different Types of Clustering, Different Types of Clusters, K-means, The Basic K-means Algorithm, K-means: Additional Issues, Bisecting Kmeans, K-means and Different Types of Clusters, Strengths and Weaknesses, K-means as an Optimization Problem, Agglomerative Hierarchical Clustering, Basic Agglomerative Hierarchical Clustering Algorithm, Specific Techniques, DBSCAN, Traditional Density: Center-Based Approach, The DBSCAN Algorithm, Strengths and Weaknesses.	8
TOTAL		50

TEXT BOOKS:

1. Introduction to Data Mining : Pang-Ning tan, Michael Steinbach, Vip Kumar, Pearson
2. Data Mining, Concepts and Techniques, 3/e Jiawei HAN ,Micheline KAMBER, Elsevier

REFERENCE BOOKS:

1. Introduction to Data Mining with Case studies 2nd ed: Gk Gupta; PH
2. Data Mining: Introductory and Advanced Topics: Dunham , Sridhar Pearson.
3. Data Warehousing .Data Mining & OLAP, Alex Berson, Stephen j Smith, TMH
4. Data Mining Theory and Practice, Somnan, Diwakar, Ajay PHI, 2006

FACULTY MEMBER

HEAD OF THE DEPARTMENT