



# SILVER OAK UNIVERSITY

## Engineering and Technology (M.Tech.)

### Computer Engineering (Software Engineering)

Subject Name: Data Mining

Subject Code:

Semester: II

#### Prerequisite:

Data Base Management System

#### Objective:

Student will be enabled to learn basics of Information Search techniques using data mining and will be able to differentiate Database Management (DBMS) and large size data warehousing methods and their needs in business systems. Student will also learn various Data mining search techniques and need of Data mart over DBMS, database and Data warehouse. Student will also learn various OLAP technologies will make student capable of handling huge information. Will be enabled to learn Data mining tools such as XML Miner and Weka and Data Mining algorithms

#### Teaching and Examination Scheme:

Teaching Scheme			Credits	Evaluation Scheme				Total Marks
L	T	P	C	Internal		External		
				Th	Pr	Th	Pr	
3	0	2	4	30	20	70	30	150

#### Content:

Sr. No.	Content	Total Hours	% Weightage
1	<b>Introduction to Data Mining</b> Importance of Data Mining, Data Mining Functionalities, Classification of Data mining systems, Data mining Architecture, Major Issues in Data Mining, Applications of Data Mining, Social impacts of data mining.	8	20
2	<b>Data Pre-processing &amp; Data Mining primitives</b> Data Pre-processing, Data cleaning, Data Integration and Transformation, Data reduction, Discretization and Concept Hierarchy Generation. Data Mining primitives, Languages and System Architectures, Concept Description: characterization and Comparison, Analytical Characterization, Mining Class Comparison.	8	20

3	<b>Association Rules &amp; Mining</b> Association Rule Mining, Mining of Single dimensional Boolean association rules, Multilevel association rules and Multidimensional association rules, Correlation analysis, Constraint based association Mining.	5	10
4	<b>Classification and Predication:</b> Basic issues regarding classification and predication, Classification by Decision Tree, Bayesian classification, and classification by back propagation, Associative classification, Prediction, Classifier Accuracy.	8	20
5	<b>Cluster Analysis:</b> Cluster Analysis, basic issues, clustering using partitioning methods, Hierarchical methods, Density based methods, Grid based methods and model based methods, Algorithms for outlier analysis.	5	15
6	<b>Mining complex Types of data:</b> Multidimensional analysis and descriptive mining of complex data objects, Introduction to spatial mining, multimedia mining, temporal mining, text mining and web mining with related algorithms.	5	15

#### Course Outcome:

Sr. No.	CO statement	Unit No
CO-1	Student will be enabled to learn basics of Data Warehouse fundamentals and will be able to differentiate Database Management (DBMS) and large size data warehousing methods and their needs in business systems.	01
CO-2	Student will also learn various Data mining search techniques and need of Data mart over DBMS, database and Data warehouse.	02
CO-3	Student will also learn various OLAP technologies will make student capable of handling huge information.	02
CO-4	Implement data mining techniques to solve problems in other disciplines in a mathematical way	04
CO-5	Will be enabled to learn Data mining tools such as XML Miner and Weka and Data Mining algorithms	06

#### Teaching & Learning Methodology:-

The concerned faculty member shall provide the list of peer reviewed Journals and Tier-I and Tier-II Conferences relating to the subject (or relating to the area of thesis for seminar) to the students in the beginning of the semester. Every student or a group of students shall critically study 2 papers, integrate the details and make presentation in the last two weeks of the semester.

**List of Experiments/Tutorials:**

- (1) Software Forensic tools such as XML Miner & WeKA
- (2) Implementation of data mining alg. such as Clustering, Classification and Association Rules
- (3) OLAP in standard SQL Server
- (4) MS Excel & Data mining plugin

**Major Equipment:**

- (1) WeKA
- (2) XML Miner

**Books Recommended:-**

1. Data Mining concepts and Techniques by Jiawei Han, Micheline Kamber –Elsevier.
2. Data Mining by Arun K. Pujari – University Press.
3. Modern Data Warehousing, Data Mining and Visualization by George M. Marakas –Pearson.
4. Data Mining by Vikram Puri and P. Radha Krishna –Oxford Press.
5. Data Warehousing by Reema Tharaja –Oxford Press