



Course Title: Data Warehousing and Mining

Course Level: UG

Course Code: CSIT321

Credit Units: 03

L	T	P/ S	SW/F W	TOTAL CREDIT UNITS
2	1	-	-	3

Course Objectives:

This course aims at

- The design and management of data warehouse (DW) and Data Mining Process.
- Giving insights on how the Data Warehouse collects and integrates data, leading to knowledge discovery.
- Introducing the core concepts of data warehousing and data mining, its techniques, implementation, benefits, and outcome expectations from this new technology.
- Data Mining (DM) process for extracting meaningful information from large volumes of data generated in an organization.
- Identifying industry branches which most benefit from DM.

Pre-requisites: Knowledge of RDBMS

Course Contents/Syllabus:

	Weightage (%)
Module I Data Warehouse fundamentals	20%
The Data Warehouse –Introduction, characteristics, its competitive advantages, Operational Database Systems and Data Warehouse (OLTP & OLAP). Multidimensional Data Models: Types of Data, from Tables and Spreadsheets to Data Cubes	
Module II Principles of dimensional modeling	20%
Identifying Facts and Dimensions, Designing Fact Tables, Designing Dimension Table, Data Warehouse Schemas, OLAP Operations, Data Extraction, Cleanup & Transformation, Star, snowflake and galaxy schemas for multidimensional databases. Architecture for a warehouse, Steps for construction of Data Warehouses, Data Marts, Metadata. Different OLAP operations, OLAP Server: ROLAP, MOLAP and HOLAP.	

Module III Data Mining	20%
From Data warehousing to data mining, Motivation, Knowledge Discovery Process, objectives of Data Mining, the business context for DM, Process improvement, marketing and CRM, Tools of Data Mining	
Module IV Data Mining Functionalities	20%
Data preparation, Data Mining Techniques: Statistical techniques, Characterization and discrimination, Association and market basket analysis, Classification and Prediction, Cluster analysis, Outlier analysis.	
Module V Data Mining Applications	20%
Text Mining, Spatial Databases, Web Mining. Case studies in building business environment. , Applications in telecommunications industry, retail, target marketing, fraud protection, health care, science, ecommerce, banking and finance.	

Student Learning Outcomes:

After completing the course, the student will be able to

- Identify the role of data mining in the providing competitive edge in business.
- Describe the components of Data Warehouse
- Model the relational database required for an enterprise data warehouse
- Extract, cleanse, consolidate, and transform heterogeneous data into a single data warehouse
- Examine the types of data to be mined.
- Analyze the data to generate information and knowledge that lead to informed decisions for businesses
- Generate insightful trends using data mining techniques.
- Select and apply proper data mining algorithms to build analytical applications

Pedagogy for Course Delivery:

The course will employ lectures, class discussions, presentations and case studies. We concentrate on demonstrating how discovering the hidden knowledge in corporate databases will help managers to make near-real time intelligent business and operation decisions. Many practical examples on realistic industry scenarios involving business intelligence systems will be discussed in the class. Through presentations, students will have a chance to evaluate a wide range of new methodologies to facilitate knowledge discovery process.

Assessment/ Examination Scheme:

Theory L/T (%)	Lab/Practical/Studio (%)	Total(%)
100%	-	100%

Theory Assessment (L&T):

Continuous Assessment/Internal Assessment					End Term Examination
Components (Drop down)	Viva	Mid Term	Case Study	Attendance	EE
Weightage (%)	5	10	10	5	70

Text books and References:**Text Books:**

- Jiawei Han & Micheline Kamber, “Data Mining: Concepts & Techniques”, Morgan Kaufmann Publishers , 2002
- Paul Raj Poonia, “Fundamentals of Data Warehousing”, John Wiley & Sons, 2004.
- Sam Anahony, “Data Warehousing in the real world: A practical guide for building decision support systems”, John Wiley, 2004

Reference Books:

- Data Ware housing: Concepts, Techniques, Products and Applications, C.S.R. Prabhu, Prentice Hall of India, 2001.
- Sam Anahory, Dennis Murray. Data Warehousing in the Real World, Pearson, 2005.
- David Taniar , Progressive methods in Data Warehousing and Business Intelligence: Concepts and competitive analytics, Idea Group Inc, 2009.