Title of Subject: DATA WAREHOUSING & DATA MINING
Disciplines: SOFTWARE ENGINEERING
Pre-requisites: Data Base Management and Administration, Data Structures and Algorithm Analysis
Assessment: 20% Sessional Marks 80% written Examination
Term: FIRST TERM FINAL YEAR
Effective: 09-Batch and onwards

Aims: This course aims to give students a good overview of the ideas and the techniques which are behind recent developments in the data warehousing and On Line Analytical Processing (OLAP) fields, in terms of data models and conceptual design methodologies and coverage of the knowledge discovery process including common data mining techniques.

Objective:
- To have an understanding of the foundations, the design, the maintenance, the evolution, and the use of data warehouses.
- To master the basic range of techniques for creating, controlling, and navigating dimensional business databases.
- To have an understanding of the data mining process, its applicability, advantages and pitfalls.

Contents:
Introduction to Data Warehousing
The Evolution of Data Warehousing; Purpose of developing Data Warehouse; Difference between Operational and Decision Support Systems; Data Warehousing; Features, benefits and applications of Data Warehouse; Data Marts & its types.

Planning and Requirements
Planning your Data warehouse; Reasons for requirements; Interviewing for Requirements; Analyzing Legacy System

Logical & Physical Data Modeling
The Database: Heart of the Data Warehouse; Database Design Life Cycle; Designing for OLTP and Designing for Decision Support; Logical Database Design; Physical Database Design.

Denormalization & Dimensional Modeling
Denormalization, Data Marts, and Data Warehouses; Denormalizing Through Prejoins; Column Replication or Movement; Pre-aggregation; Dimensional Modeling: Star Schema, Snowflake Schema, and Fact Constellation Schema.
Data Extraction, Transformation and Loading
ETL (Extract, Transform, Load) Process overview; ETL versus ELT; Data Extraction, Transformation, Cleansing, Integration and Staging.

Online Analytical Processing (OLAP) Implementation Techniques
Demand for OLAP; Major Features and Functions: Cubes, Hyper-cubes, Drill-Down and Roll-Up, Slice and Dice; OLAP Models: MOLAP, ROLAP, And HOLAP.

Data Warehousing and the Web
Web-enabled Data Warehouse, Web-enabled Information Delivery, New Information Strategies, Browser Technology for the Data Warehouse, Security Issues

Data Mining
Introduction to Data Mining; The Knowledge Discovery process; OLAP versus Data Mining; Data Mining & Data Warehouse; Applications and trends in Data Mining; Major Data Mining Techniques: Cluster Detection, Decision Trees, Memory-Based Reasoning, Link Analysis, Genetic Algorithms & Neural Networks; Data Mining Tools; Social Network Mining.

Note: Practical will be based on theory

Recommended books:
3. Jiawei Han & Kamber M., Data Mining: Concepts & Techniques, 2nd ed, 2006, Morgan Kaufman Publisher
4. NCR Teradata University Program Guide

Approval: 
Board of Studies 1st Meeting Res.No 01.00 Dated:18.11.09
Board of Faculty of EEC Engg: 5th Meeting Res.No…5.6… Dated: 18.12.09
Academic Council 73rd Meeting Res.No…16 Dated: 23.12.09