Course Title: Data Warehousing and Data Mining
Course no: CSC- 451   Full Marks**:** 60+20+20
Credit hours: 3   Pass Marks**:** 24+8+8

Nature of course: Theory (3 Hrs.) + Lab (3 Hrs.)

Course Synopsis: Analysis of advanced aspect of data warehousing and data mining.

Goal: This course introduces advanced aspects of data warehousing and data mining, encompassing the principles, research results and commercial application of the current technologies

Course Contents:

Unit- 1   **5 Hrs.**

Concepts of Data Warehouse and Data Mining including its functionalities, stages of Knowledge discovery in database(KDD) , Setting up a KDD environment, Issues in Data Warehouse and Data Mining, Application of Data Warehouse and Data Mining

Unit-2   **4 Hrs.**

DBMS vs. Data Warehouse, Data marts, Metadata, Multidimensional data model, Data Cubes, Schemas for Multidimensional Database: Stars, Snowflakes and Fact Constellations.

Unit- 3   **6 Hrs.**

Data Warehouse Architecture, Distributed and Virtual Data Warehouse, Data Warehouse Manager, OLTP, OLAP, MOLAP, HOLAP, types of OLAP, servers.

Unit- 4   **4 Hrs.**

Computation of Data Cubes, modeling: OLAP data, OLAP queries, Data Warehouse back end tools, tuning and testing of Data Warehouse.

Unit- 5   **4Hrs.**

Data Mining definition and Task, KDD versus Data Mining, Data Mining techniques, tools and application.

Unit- 6   **5Hrs.**

Data mining query languages, data specification, specifying knowledge, hierarchy specification, pattern presentation & visualization specification, data mining languages and standardization of data mining.

Unit- 7   **6 Hrs.**

Mining Association Rules in Large Databases: Association Rule Mining, why Association Mining is necessary, Pros and Cons of Association Rules, Apriori Algorithm.

Unit- 8   **7 Hrs.**

Classification and Prediction: Issues Regarding Classification and Prediction, Classification by Decision Tree Induction, Introduction to Regression, Types of Regression, Introduction to clustering, K-mean and K-Mediod Algorithms.

Unit- 9   **4 Hrs.**

Mining Complex Types of Data: Mining Text Databases, Mining the World Wide Web, Mining Multimedia and Spatial Databases.

Laboratory Works: Cover all the concept of datawarehouse and mining mention in a course

Samples
1. Creating a simple data warehouse
2. OLAP operations: Roll Up, Drill Down, Slice, Dice through SQL- Server
3. Concepts of data cleaning and preparing for operation
4. Association rule mining though data mining tools
5. Data Classification through data mining tools
6. Clustering through data mining tools
7. Data visualization through data mining tools

Reference books:

1. Data Mining Concepts and Techniques, Morgan Kaufmann J. Han, M Kamber Second Edition ISBN: 978-1-55860-901-3
2. Data Warehousing in the Real World – Sam Anahory and Dennis Murray, Pearson Edition Asia.
3. Data Mining Techniques – Arun K Pujari, University Press.
4. Data Mining- Pieter Adriaans, DolfZantinge
5. Data Mining, Alex Berson,StephenSmith,KorthTheorling,TMH.
6. Data Mining, Adriaans, Addison-Wesley Longman.