CS6L3   **DATA MINING LAB**   **Credits: 2**
(Common to CSE & IT)
Required

Lecture: --   **Internal assessment: 25 marks**
Lab: 3 periods/week   **Semester end examination: 50 marks**

Coruse context and Overview: The data mining process includes data selection and cleaning, machine learning techniques to ``learn" knowledge that is ``hidden" in data, and the reporting and visualization of the resulting knowledge. This course will cover these issues and will illustrate the whole process by examples of practical applications from the life sciences, computer science, and commerce.

Prerequisites: SQL, DBMS

_____

Objective:

To implement metadata import agents to pull metadata from leading business intelligence tools and populate a metadata repository.

1. Import metadata from specific business intelligence tools and populate a Meta data repository.
2. Load data from heterogeneous sources including text files into a pre-defined warehouse schema. Case Study
3. Design a data mart from scratch to store the credit history of customers of a bank. Use this credit profiling to process future loan applications.
4. Design and build a Data Warehouse using bottom up approach taking any relevant example.

Learning Outcomes:
>   Ability to:

1. Analyze and interpret data using Weka Tool.
2. Apply frequent item set generation algorithms.
3. Apply classification and clustering algorithms.

WEKA Tool is used for the following experiments

Data Warehousing:

Objective: To implement metadata import agents to pull metadata from leading businessintelligence tools and populate a metadata repository.

1. Import metadata from specific business intelligence tools and populate a Meta data repository.

2. Load data from heterogeneous sources including text files into a pre-defined warehouse schema. Case Study

3. Design a data mart from scratch to store the credit history of customers of a bank. Use this credit profiling to process future loan applications.

4. Design and build a Data Warehouse using bottom up approach taking any relevant example.

Data Mining:

Objective: Using various Data Mining techniques understand Data Mining processes.

LIST OF PROGRAMS

1. **Introduction to Data Mining tool WEKA**
2. **Creating Arff dataset.**
3. **Preprocessing techniques on dataset.**
4. **Generating frequent itemset using Apriori Algorithm.**
5. **Generating frequent itemset using FP-Growth Algorithm.**
6. **Classifying the data samples using Decision Tree Induction.**
7. **Classifying the given data samples using Naive Bayes technique.**
8. **Cluster analysis using K-Means algorithm.**
9. **Cluster Analysis Hierarchical Algorithm.**
10. **Cluster Analysis using Density Based Clustering Algorithm**

Learning Resources

Reference Books:

1.Data Mining : Methods and Techniques , A B M Shawkat Ali, Saleh A. Wasimi , Cengage Learning.

2. The Data Warehouse Life Cycle Tool kit , RALPH KIMBALL, Wiley Student Edition.