UNIVERSITY OF MADRAS

B.Sc. DEGREE PROGRAMME IN COMPUTER SCIENCE

SYLLABUS WITH EFFECT FROM 2023-2024

Year: III **Semester: V**

Data Mining and Warehousing Common for B.C.A. , B.Sc.-SA , B.Sc.-CSc , B.Sc.-CSc-wAI , B.Sc.-CSc-wDS	**325E5C**
Credits 3	**Lecture Hours: 4 per week**

Learning Objectives: (for teachers: what they have to do in the class/lab/field)

- To provide the knowledge on Data Mining and Warehousing concepts and techniques.
- To study the basic concepts of cluster analysis
- To study a set of typical clustering methodologies, algorithms, and applications

Course Outcomes: (for students: To know what they are going to learn)

CO1: To understand the basic concepts and the functionality of the various data mining and data warehousing component

CO2: To know the concepts of Data mining system architectures

CO3: To analyse the principles of association rules

CO4: To get analytical idea on Classification and prediction methods.

CO5: To Gain knowledge on Cluster analysis and its methods.

Units	Contents
I	Introduction: Data mining – Functionalities – Classification – Introduction to Data Warehousing – Data Pre-processing: Pre-processing the Data – Data cleaning – Data Integration and Transformation – Data Reduction
II	Data Mining, Primitives, Languages and System Architecture: Data Mining – Primitives – Data Mining Query Language, Architecture of Data mining Systems. Concept Description, Characterization and Comparison: Concept Description, Data Generalization and Summarization, Analytical Characterization, Mining Class Comparison – Statistical Measures
III	Mining Association Rules: Basic Concepts – Single Dimensional Boolean Association Rules From Transaction Databases, Multilevel Association Rules from transaction databases – Multi dimension Association Rules from Relational Database and Data Warehouses
IV	Classification and Prediction: Introduction – Issues – Decision Tree Induction – Bayesian Classification – Classification of Back Propagation. Classification based on ConceptsfromAssociationRuleMining–OtherMethods. Prediction – Introduction – Classifier Accuracy.
V	Cluster Analysis: Introduction – Types of Data in Cluster Analysis, Petitioning Methods – Hierarchical Methods-Density Based Methods – GRID Based Method – Model based Clustering Method

Learning Resources:

Recommended Texts

1. Han and M. Kamber, "Data Mining Concepts and Techniques", 2001, Harcourt India Pvt. Ltd, New Delhi.

Reference Books

1. K.P. Soman, Shyam Diwakar, V. Ajay "Insight into Data Mining Theory and Practice ",Prentice Hall of India Pvt. Ltd, New Delhi
2. Parteek Bhatia, 'Data Mining and Data Warehousing: Principles and Practical Techniques',Cambridge University Press, 2019

Web resources: Web resources from NDL Library, E-content from open-source libraries