UNIVERSITY OF MADRAS

M.Sc. DEGREE PROGRAMME IN COMPUTER SCIENCE

SYLLABUS WITH EFFECT FROM 2023-2024

Title of the Paper	**Data Mining and Warehousing**		
Core–V - Theory	I Year & II Semester	Credit: 4	436C2A

Objectives:

- Enable the students to learn the concepts of Mining tasks, classification, clustering and Data Warehousing.
- Develop skills using recent datamining software for solving practical problems.
- Develop and apply critical thinking, problem-solving, and decision-making skills.

Outcomes:

1.	Understand the basic datamining techniques and algorithms	K1, K2
2.	Understand the Association rules, Clustering techniques and Data warehousing contents	K2, K3
3.	Compare and evaluate different datamining techniques like classification, prediction, Clustering and association rule mining	K4, K5
4.	Design data warehouse with dimensional modelling and apply OLAP operations	K5, K6
5.	Identify appropriate datamining algorithms to solve real world problems	K6
K1-Remember;K2-Understand;K3-Apply;K4-Analyze;K5-Evaluate; K6-Create		

UNIT I: Basic data mining tasks – data mining versus knowledge discovery in databases – data mining issues – data mining metrics – social implications of data mining – data mining from a database perspective. Data mining techniques: Introduction – a statistical perspective on data mining – similarity measures – decision trees – neural networks – genetic algorithms.

UNIT II: Classification: Introduction –Statistical –based algorithms -distance–based algorithms - decision tree-based algorithms -neural network-based algorithms– rule-based algorithms– combining techniques.

UNIT III: Clustering: Introduction – Similarity and Distance Measures–Outliers–Hierarchical Algorithms - Partitional Algorithms. Association rules: Introduction - large item sets - basic algorithms – parallel &distributed algorithms – comparing approaches- incremental rules – advanced association rules techniques – measuring the quality of rules.

UNIT IV: Data warehousing: introduction – characteristics of a data ware house–data marts– other aspects of data mart. Online analytical processing: introduction -OLTP & OLAP systems Data modelling – star schema for multidimensional view –data modelling – multi fact star schema or snow flake schema – OLAP TOOLS – State of the market – OLAP TOOLS and the internet.

UNIT V: Developing a data WAREHOUSE: why and how to build a data warehouse –data warehouse architectural strategies and organization issues - design consideration – data content – metadata distribution of data – tools for data warehousing – performance considerations – crucial decisions in designing a data warehouse. Applications of data warehousing and data mining in government: Introduction - national data warehouses – other areas for data warehousing and data mining.

Text Books

1. Margaret H. Dunham, "Data Mining: Introductory and Advanced Topics", Pearson education,2003.
2. C.S.R. Prabhu, "Data Warehousing Concepts, Techniques, Products and Applications", PHI, Second Edition.

Reference Books

1. Arun K. Pujari, "Data Mining Techniques", Universities Press (India) Pvt. Ltd.,2003.
2. Alex Berson, Stephen J. Smith, "Data Warehousing, Data Mining and OLAP", TMCH, 2001.
3. Jiawei Han &Micheline Kamber, Academic press. "Data Mining Concepts & Techniques", 2001,

Related Online Contents

1. https://www.javatpoint.com/data-warehouse
2. https://nptel.ac.in/noc/courses/noc20/SEM1/noc20-cs12/
3. https://www.btechguru.com/training--it--database-management-systems--file-structures--introduction-to-data-warehousing-and-olap-2-video-lecture--12054--26--151.html

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO1	S	M	S	S	S	S	M	M	M	M
CO2	S	S	S	S	S	S	S	M	S	S
CO3	S	S	S	S	S	S	S	M	S	S
CO4	S	S	S	S	S	S	S	M	S	S
CO5	S	S	S	S	S	S	S	M	S	S

*S-Strong; M-Medium; L-Low