

MCAS4110: DATA WAREHOUSING AND MINING

AIM: To studies the basic principles of data mining and data warehousing architecture.

Unit-I

11 periods

Data Mining: Introduction – Information and production factor – Data mining Vs Query tools – Data and machine learning- Machine learning and statistics-Data Mining in marketing – Data Mining and ethics- Nuggets and data mining- Database Mining – A performance and database Perspective- Self learning computer systems – Concept learning – Data mining and the Data Warehousing

Unit-II

13 periods

Knowledge Discovery Process : Knowledge discovery process – Data selection – Cleaning – Enrichment – Coding – Preliminary analysis of the data set using traditional query tools – Visualization techniques – Knowledge representation- Decision trees – Classification rules- Association rules –Rules with exceptions- rules involving relations- Trees for numeric - Instance-based representation- Neural Networks – Genetic Algorithms – Clustering - KDD (Knowledge Discovery in Databases) Environment.

Unit-III

13 periods

Dataware House – Architecture: Data warehouse Architecture – System Process – Process Architecture – Design – Database Schema – Partitioning Strategy – Aggregations – Data Marting – Meta Data – System and Data Warehouse Process Managers.

Unit-IV

12 periods

Hardware and Operational Design: Hardware and operational design of Data Warehouse – Hardware Architecture – Physical Layout – Security – Backup and Recovery – Service – Level Agreement – Operating the Warehouse.

Unit-V

11 periods

Planning- Tuning and Testing: Capacity planning – Tuning the Data Warehouse – Testing Warehouses – Data Warehouse Features.

Text Books:

1. Pieter Adriaans, Dolf zantinge, “Data Mining”, Pearson Education, 2007.
1. Sam Anahory, Dennis Murray, “Data Warehousing in the real world – A Practical Guide for Building Decision Support Systems”, Pearson Education, 2006.

Reference Books:

1. Ian.H.Witten & Eibe Frank, “Data Mining – Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers, 2006.
2. Jiawei Han, Micheline Kamber, “Data Mining: Concepts and Techniques” Morgan Kaufmann Publishers, 2000.
3. Hanand J and M. Kamber, “Data Mining: Concepts and Techniques”, Second Edition, Morgan Kaufman, 2006.