

CORE COURSE III - DISTRIBUTED DATABASE SYSTEMS

UNIT I

INTRODUCTION:- Major issues in data Mining – What is Data Warehouse – Difference between Database Systems and Data Warehouse Data Warehouse Architecture – Data Warehouse Implementation.

Unit II

Data Pre processing : Data Cleaning – Integration – Transformation – Reduction Data Mining Primitives – Data Mining Query Language – Architecture of Data Mining System

Unit III

Association rules in Large Databases: Mining single and multilevel association rules from Transactional Database, Relational databases and Data Warehouses.

Unit IV

Classification : Issues – Decision Tree Induction – Baseline Classification – Classification by propagations – other Classification methods – Prediction – Classifier Accuracy.

Unit V

Cluster Analysis – Types – Categorization – Hierarchical methods – Density based method – Grid based method – Model based method – Outlier Analysis Applications: Data Mining Applications – Themes on Data Mining – Commercial DM

Text Book:

1. Data Mining : “Concepts and Techniques”, Jiawei Han, Micheline Kamber, Morgan Kaufmann Publishers, 2001 by Academic Press.

Reference Books:

1. Usama M. Fayyad, Gregory Piatetsky – Shapiro, Padhraí Smyth, Ramasamy Uthurusamy, “Advances in Knowledge Discovery and Data mining”. The M.I.T. Press, 1996.
2. Arun K. Pujari, “Data Mining Techniques”, University Press India Ltd, 2001