Code No.: MCA 204 **L T C**
Paper: Data Warehousing & Data Mining **3 1 4**

INSTRUCTIONS TO PAPER SETTERS:

1. Question No. 1 should be compulsory and cover the entire syllabus. There should be 10 questions of short answer type of 2 marks each, having at least 2 questions from each unit.
2. Apart from Question No. 1, rest of the paper shall consist of four units as per the syllabus. Every unit should have two questions to evaluate analytical/technical skills of candidate. However, student may be asked to attempt only 1 question from each unit. Each question should be 10 marks including subparts, if any.

OBJECTIVES: *This course is an attempt to provide you with the basic information about data ware house and their development. This course also provides the basic conceptual background necessary to design and develop data ware house applications.*

PRE-REQUISITE:

- Information System Concepts

UNIT- I

The Compelling Need for data warehousing: Escalating Need for strategic information, failures of Past decision-support systems, operational versus decision-support systems, data warehousing – the only viable solution, data warehouse defined Data warehouse – The building Blocks: Defining Features, data warehouses and data marts, overview of the components, metadata in the data warehouse Defining the business requirements: Dimensional analysis, OLAP operations : Drill-down and roll-up, slice-and-dice or rotation.

[No. of Hrs: 11]

UNIT- II

Principles of dimensional modeling: , the STAR schema, STAR Schema Keys, Advantages of the STAR Schema Dimensional Modeling: Updates to the Dimension tables, miscellaneous dimensions, the snowflake schema, aggregate fact tables, families of STARS ,Steps for the Design & Construction of Data warehouse : Framework , Architecture , Type of OLAP Servers : ROLAP , MOLAP , Data warehouse implementation tolls & techniques.

[No. of Hrs.: 10]

UNIT- III

Data Mining, Data Mining of what kind of Data , Knowledge discovery process (KDD) , What kind of patterns can be mined , OLAP versus data mining, data mining and the data warehouse, Data mining functionalities, classification Systems , Data processing : Cleaning , Integration & transformation, Reduction . Data Mining primitives: What defines a Data Mining Task.

[No. of Hrs.: 10]

UNIT- IV

Data Mining Query language (DMQL), Cluster Analysis : Partitioning , Hierarchical Density , Grid & Model based methods ., Major Data Mining Techniques, Cluster detection, decision trees, memory-based reasoning, link analysis, neural networks, genetic algorithms, moving into data mining, Data Mining Applications, Benefits of data mining & applications.

[No. of Hrs.: 11]

TEXT BOOKS:

1. Paul Raj Poonia, "Fundamentals of Data Warehousing", John Wiley & Sons, 2003.
2. W. H. Inmon, "Building the Operational Data Store",2nd Ed., John Wiley, 1999
3. Sam Anahony, "Data Warehousing in the Real World: A Practical Guide for Building Decision Support Systems", John Wiley, 2004.
4. Jarke, "Fundamentals of Data Warehouse", Springer

REFERENCES:

1. Kamber and Han, "Data Mining Concepts and Techniques", Hartcourt India P. Ltd., 2001.
2. G. K. Gupta, "Introduction to Data Mining with Case Studies", PHI, 2006.
3. A. B. M. Shawkat Ali, Saleh A. Wasimi, "Data Mining Methods and Techniques", Cengage Learning, 2009.
4. Pang - Ning, Michael- Steinbach, "Introduction to Data Mining", Pearson, 4th Ed., 2009.