

**Code No.: ETCS 310**

**Paper: Data Warehousing and Data Mining**

<b>L</b>	<b>T</b>	<b>C</b>
<b>3</b>	<b>1</b>	<b>4</b>

**INSTRUCTIONS TO PAPER SETTERS:**

**MAXIMUM MARKS: 75**

1. Question No. 1 should be compulsory and cover the entire syllabus. This question should have objective or short answer type questions. It should be of 25 marks.
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. However, student may be asked to attempt only 1 question from each unit. Each question should be of 12.5 marks.

### **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, information packages – a new concept, requirements gathering methods, requirements definition: scope and content

**[No. of Hrs.: 11]**

### **UNIT – II**

Principles of dimensional modeling: Objectives, From Requirements to data design, 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

**[No. of Hrs.: 11]**

### **UNIT – III**

OLAP in the Data Warehouse: Demand for Online analytical processing, need for multidimensional analysis, fast access and powerful calculations, limitations of other analysis methods, OLAP is the answer, OLAP definitions and rules, OLAP characteristics, major features and functions, general features, dimensional analysis, what are hypercubes? Drill-down and roll-up, slice-and-dice or rotation, OLAP models, overview of variations, the MOLAP model, the ROLAP model, ROLAP versus MOLAP, OLAP implementation considerations

**[No. of Hrs.: 11]**

### **UNIT – IV**

Data Mining Basics: What is Data Mining, Data Mining Defined, The knowledge discovery process, OLAP versus data mining, data mining and the data warehouse, 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 in retail industry, applications in telecommunications industry, applications in banking and finance.

**[No. of Hrs.: 11]**

### **TEXT BOOKS:**

1. Paul Raj Poonia, “Fundamentals of Data Warehousing”, John Wiley & Sons, 2004.
2. Sam Anahony, “Data Warehousing in the real world: A practical guide for building decision support systems”, John Wiley, 2004

### **REFERENCES BOOKS:**

1. W. H. Inmon, “Building the operational data store”, 2<sup>nd</sup> Ed., John Wiley, 1999.
2. Kamber and Han, “Data Mining Concepts and Techniques”, Hartcourt India P. Ltd., 2001

- 
- 
3. "Data Warehousing", BPB Publications, 2004.