Paper Code: BCA 302 | **L** | **T** | **C**
Paper ID: 20302 | **3** | **1** | **4**
Paper: Data warehouse and data mining
Pre-requisites:
- **Information System Concepts**

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.

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 12.5 marks

UNIT – I

Data mining: Inttroduction, Data mining – on what kind of data, data mining functionalities – what kind of patterns to be mined, Classification of data mining systems, data mining task primitives, integration of a data mining systems with a database or data warehouse systems, major issues in data mining.

Data preprocessing: Descriptive data summarization, data cleaning, data integration and transformation, data reduction, data descretization and concept hierarchy generation.

[No. of Hrs: 11]

UNIT – II

Data warehouse and OLAP technology: What is data warehouse, A multidimensional data model, data warehouse architecture, data warehouse implementation, data warehouse usage, OLAP, OLAM

Mining frequent patterns, association and correlation, efficient and scalable frequent itemset mining methods, From association mining to correlation analysis.

[No. of Hrs: 11]

UNIT – III

Classification and prediction: Introduction, issues, classification by decision tree induction, rule based classification, classification by back propagation, lazy learners, other classification methods, Prediction: accuracy and error measures, evaluating the accuracy of a classifier or predictor.

Cluster Analysis: Types of data in cluster analysis, a categorization of major clustering methods, partitioning methods.

[No. of Hrs: 11]

UNIT – IV

Mining complex types of data: Multidimensional analysis and descriptive mining of complex data objects, mining spatial database, multimedia database, mining world wide web.

Applications and trends in data mining: Data mining applications, data mining system products and research prototypes, social impact of data mining, trends in data mining.

[No. of Hrs: 11]

TEXT BOOKS:

[T1] Kamber and Han, "Data Mining Concepts and Techniques", Hartcourt India P. Ltd.,2001.
[T2] Paul Raj Poonia, "Fundamentals of Data Warehousing", John Wiley & Sons, 2003.