Paper code: ARM 206	L	P	Credits
Subject: Data Warehousing and Data Mining	3	0	3

Marking Scheme:
Teachers Continuous Evaluation: As per university examination norms from time to time.
End Term Theory Examination: As per university examination norms from time to time.

INSTRUCTIONS TO PAPER SETTERS: Maximum Marks : AS per University norms

- There should be 9 questions in the end term examination question paper
- Question No. 1 should be compulsory and cover the entire syllabus. This question should have objective or short answer type questions.
- Apart from Question No. 1, the rest of the paper shall consist of four units as per the syllabus. Every unit should have two questions. However, students may be asked to attempt only 1 question from each unit.
- The questions are to be framed keeping in view the learning outcomes of course/paper. The standard/ level of the questions to be asked should be at the level of the prescribed textbooks.
- The requirement of (scientific) calculators/ log-tables/ data-tables may be specified if required

CO1:	Ability of students to understand the concepts of Data warehouse and OLAP **[K1,K2]**
CO2:	Ability of students to explore the basic concepts of Data Mining **[K2]**
CO3:	Ability of students to perform Data Mining using Regression, Classification and Clustering. **[K3]**
CO4:	Ability of students to demonstrate applications of Data Mining**[K3]**

CO/PO	**PO01**	**PO02**	**PO03**	**PO04**	**PO05**	**PO06**	**PO07**	**PO08**	**PO09**	**PO10**	**PO11**	**PO12**
CO1	3	3	3	3	3	2	1	1	2	1	1	2
CO2	2	3	3	3	3	2	1	1	1	1	1	2
CO3	2	3	3	3	3	2	1	1	2	1	2	3
CO4	3	3	3	3	3	2	1	1	1	1	2	3

Course Content	**No of lectures**
Unit I Data Warehousing and Business Analysis: - Data warehousing Components, Building a Data warehouse, Data Warehouse Architecture, DWH Schemas for Decision Support, Data Extraction-Cleaning-Transformation Tools, Metadata, Reporting, Query tools and Applications, Online Analytical Processing (OLAP), OLAP and Multidimensional Data Analysis, OLAP vs OLTP	**[10]**
Unit II	**[10]**

Data Mining: - Data Mining Fundamentals, Data Pre-processing, Data Cleaning, Data Integration and Transformation, Data Reduction, Architecture of a typical Data Mining systems, Classification of Data Mining Systems. Association Rule Mining: - Frequent Itemset Mining Methods, Mining Various Kinds of Association Rules, Association Mining to Correlation Analysis, Apriori algorithm	
Unit III Classification and Prediction: - Issues Regarding Classification and Prediction, Regression-Single variate and multivariate, Accuracy and Error, Classification by Decision Tree Induction, Bayesian Classification, Rule Based Classification, Classification by Back propagation, Support Vector Machines, Lazy Learners, Evaluating the Accuracy of a Classifier or Predictor, efficient Model Section.	**[10]**
Unit IV Cluster Analysis: - Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical methods, Density-Based Methods, Clustering High-Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis. Applications of Data Mining: Multimedia Data Mining, Text Mining, Mining the World Wide Web.	**[10]**
Text Books: [T1] Paulraj Ponnia, "Fundamentals of Data Warehousing", John Wiley & Sons, 2004. [T2] Kamber and Han, "Data Mining Concepts and Techniques", Hart Court India P. Ltd. Elsevier Publications Second Edition, 2001	
Reference Books: [T1] Alex Berson and Stephen J. Smith, ―Data Warehousing, Data Mining & OLAP‖, Tata McGraw – Hill Edition, 35th Reprint 2016. [T2] K.P. Soman, Shyam Diwakar and V. Ajay, ―Insight into Data Mining Theory and Practice, Eastern Economy Edition, Prentice Hall of India, 2006.	