END TERM EXAMINATION

SIXTH SEMESTER [BCA] MAY-2014

Paper Code: BCA-302	**Subject: Data Warehousing and Data Mining**
	(Batch 2011 Onwards)
Time : 3 Hours	**Maximum Marks: 75**

Note: Attempt any five question, including Q.no.1 which is compulsory.

Q1 Answer the following questions in brief:-   (2.5x10=25)
 (a) What is metadata? Why it is important?
 (b) Discuss different types of data on which data mining can be applied.
 (c) What do you mean by the term 'Hierarchy'? Discuss.
 (d) Discuss the term support and confidence used in association rule mining.
 (e) Define the terms: Discrimination and characterization.
 (f) Discuss different OLAP operations.
 (g) What do you understand by outlier analysis?
 (h) What is difference between data warehouse and data mart? Discuss.
 (i) What is difference between data, information and knowledge?
 (j) What do you understand by back propagation? Discuss.

Q2 (a) What are the features of a data warehouse? Explain each feature by
        taking an example.   (6)
   (b) Discuss the architecture for data warehouse?   (6.5)

Q3 Differentiate between the following:-   (12.5)
 (a) OLAP and OLTP.
 (b) OLAP and data mining.
 (c) Classification and Prediction.

Q4 (a) Discuss the benefits of data mining.   (6)
   (b) What are the essentials steps of data mining?   (6.5)

Q5 (a) What are the applications of data mining in marketing?   (6)
   (b) Discuss the term 'data cleaning'. What are different methods for
       handling the missing values?   (6.5)

Q6 (a) Discuss the steps of decision tree classification.   (6)
   (b) Compare the advantages and disadvantages of eager classification
       e.g., decision tree, neural network) versus lazy classification (e.g., k-
       nearest neighbor, case-based reasoning).   (6.5)

Q7 What is cluster analysis? Discuss major clustering methods. (12.5)

Q8 Write note on the following:-   (12.5)
 (a) Data Transformation
 (b) Supervised learning vs. unsupervised learning.
 (c) Correlation analysis

****

P