END TERM EXAMINATION

SIXTH SEMESTER [BCA] MAY-2008

Paper Code: BCA312   Subject: Data Ware Housing And Data Mining
Paper Id: 20312   (Batch: 2001-2004)

Time : 3 Hours   Maximum Marks :75

Note: Attempt five questions in all including Q.1 which is compulsory.

Q1   Attempt **any five** parts:-   (5x3=15)
- (a) Differentiate data base and data ware housing. How DW concepts make it superior to conventional DBMS.
- (b) Differentiate legacy data base and operational database.
- (c) What is data mining? How is it different from conventional query approach in DBMS?
- (d) Describe OLAP tools. What is ROLAP?
- (e) Cluster analysis and its application in data mining.
- (f) Mining associated rules, their concept and applications.
- (g) Data mining is also termed as Knowledge Discovery from Data (KDD).
- (h) What are data mining primitives? Discuss their applications in brief.
- (i) Describe importance of initial data cleaning for use in data mining.

Q2
- (a) Discuss concept of data mart.
- (b) Describe the development of data cube technology. How does this technology make mining process more efficient?   (5+10=15)

Q3
- (a) Describe data integration and data reduction.
- (b) What are three phases of a decision tree? Describe the significance of each of the phases.   (8+7=15)

Q4
- (a) What is meta data? Describe various types used in DW and for what purpose?
- (b) How are data mining techniques useful as DSS? Discuss with examples.   (7+8=15)

Q5   To establish system for Datamining is, if necessary, to first set up a data warehousing. Discuss. Also, write various steps for preprocessing of data for application of datamining tools.   (15)

Q6   Write in brief on the following aspects of data mining:-   (5x3=15)
- (a) Spatial data bases
- (b) Temporal (time sequence) data basis
- (c) Multimedia data basis
- (d) Documents
- (e) WWW

Q7   Write briefly about:-
- (a) Statistical tools in datamining techniques.
- (b) Datamining and query language.
- (c) Datamining and privacy.   (3x5=15)

Q8   Write short notes on **any three**:-   (3x5=15)
- (a) Outlier analysis
- (b) Issues related to classification and prediction
- (c) Datamining and CRM (Customer Relation Management)
- (d) Datamining and visualization
- (e) Trends in research in datamining and its applications.

****

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

****

END TERM EXAMINATION

SIXTH SEMESTER [BCA] MAY-JUNE-2015

Paper Code: BCA-302   Subject: *Data Warehouse and Data Mining*
(Batch 2011 onwards)

Time: 3 Hours   Maximum Marks: 75

Note: Attempt any five questions including Question No. 1 which is compulsory.

Q1. Answer the following questions in brief:-   (2.5x10=25)
  (a) What is data warehousing and why it is important for decision support?
  (b) What is data preprocessing? Discuss.
  (c) Explain different OLAP operations.
  (d) What do you understand by frequent patterns? Explain.
  (e) Data mining is multidisciplinary field. Discuss.
  (f) Differentiate between Classification and Clustering.
  (g) What are the applications of Data mining? Discuss.
  (h) Define support and confidence.
  (i) Can a data mining system generate interesting patterns? Justify.
  (j) What is knowledge discovery in Database? Explain.

Q2 (a) Discuss Data integration and Data transformation in detail.   (6)
  (b) Explain the architecture of a data mining system with suitable
      diagram.   (6.5)

Q3 Differentiate between the following:-
  (a) OLAP and OLAM   (4)
  (b) Data mining and KDD   (4)
  (c) Classification and Prediction.   (4.5)

Q4 (a) What is back propogation? How is back propogation performed?   (6)
  (b) Explain decision tree in classification.   (6.5)

Q5 (a) What are the major issues in data mining? Discuss.   (6)
  (b) Discuss the social impact of Data mining detail.   (6.5)

Q6 (a) How would you measure the quality of Cluster? Explain.   (6)
  (b) List the different categorization of OLAP tools.   (6.5)

Q7 (a) Explain the design and construction of Data warehouse in detail. (12.5)

Q8 Write note on **any two** the following:-   (2X6.25=12.5)
  (a) Temporal database vs Sequential database.
  (b) Classification vs Prediction.
  (c) Lazy learners.

****

P

(Please write your Exam Roll No.)   Exam Roll No. ....

END TERM EXAMINATION

Sixth Semester [BCA] May 2017

Paper Code: BCA-302	Subject: Data warehouse and Data Mining
Time: 3 Hours	Maximum Marks: 75

Note: Attempt any five questions including Q no.1 which is compulsory.
Select one question from each unit.

Q1 Answer **any five** of the following questions briefly:-   (5x5=25)
 (a) How to classify various data mining systems?
 (b) What do you mean by Association Rule Mining?
 (c) What are data mining task primitives?
 (d) Briefly explain OLAP.
 (e) How to evaluate the accuracy of a classifier?
 (f) What are important social impacts of data mining?

UNIT-I

Q2 What are the major issues in Data Mining? Explain in detail.   (12.5)

Q3 What is data preprocessing? Write and explain all the steps of data preprocessing.   (12.5)

UNIT-II

Q4 What is Multidimensional Data Model? Explain in detail the architecture of data warehouse.   (12.5)

Q5 What are frequent patterns, association and correlation? Explain with the help of a suitable example the advantages of frequent pattern mining.   (12.5)

UNIT-III

Q6 What is Classification? Write and explain the decision tree induction method.   (12.5)

Q7 What do you understand by Cluster Analysis? Explain Partitioning method in detail.   (12.5)

UNIT-IV

Q8 Write the major applications of Data mining in today's world.   (12.5)

Q9 Write the steps to mine spatial and multimedia databases.   (12.5)

****

P

END TERM EXAMINATION

SIXTH SEMESTER [BCA] MAY 2018

Paper Code: BCA-302	Subject: Data Warehouse and Data Mining
Time: 3 Hours	Maximum Marks: 75

Note: Attempt any five questions including Q.no.1 which is compulsory.
Select one question from each unit.

Q1   Answer the following questions in brief: (2.5x10=25)
  (a) What do you mean by Frequent Pattern mining in databases?
  (b) Define Quantile plot and scatter plot.
  (c) What do you mean by data transformation? Explain.
  (d) Define Regression technique for predictive analysis of data.
  (e) Explain support and confidence in relation to association rule mining.
  (f) What is a data warehouse?
  (g) What is Starnet Query Model for querying multidimensional databases?
  (h) What are the applications of data mining?
  (i) What is accuracy and error measure in relation to classification?
  (j) What are data mining task primitives?

Unit-I

Q2   (a) Discuss the benefits of data mining.   (6.5)
  (b) How can you measure dispersion of data? Explain the concept of Range, Quartiles, Outliers, and Boxplots.   (6)

Q3   (a) How are missing values and noisy data handled in data cleaning step of data mining?   (6.5)
  (b) Define Knowledge Discovery in Databases with suitable diagram.   (6)

Unit-II

Q4   (a) Explain star, snowflake schema for multidimensional data models.(6.5)
  (b) Compare OLAP and OLTP systems.   (6)

Q5   (a) What are typical OLAP operations? Explain in brief.   (6.5)
  (b) Explain Apriori Algorithm of association rule mining.   (6)

Unit-III

Q6   (a) Explain decision tree in classification.   (6.5)
  (b) Explain k-means algorithm in detail with suitable diagram.   (6)

Q7   (a) Differentiate between eager and lazy learners? Explain their different types.   (6.5)
  (b) Explain k-medoid algorithm in detail with suitable diagram.   (6)

Unit-IV

Q8   (a) What are the major issues in data mining? Explain.   (6.5)
  (b) What are data mining applications in Telecom industry?   (6)

Q9   (a) Discuss social impacts of data mining in detail.   (6.5)
  (b) Write short notes on: (6)
    (i) Mining the World Wide Web
    (ii) Mining spatial databases.

****

P

END TERM EXAMINATION

SIXTH SEMESTER [BCA] MAY 2019

Paper Code: BCA-302	Subject: Data Warehouse and Data Mining
Time: 3 Hours	Maximum Marks: 75

Note: Attempt any five questions including Q.no.1 which is compulsory.
Select one question from each unit.

Q1   Answer the following questions in brief (Any Ten):   (2.5x10=25)

    (a) Explain the steps involved in data mining when viewed as a process of knowledge discovery.
    (b) What is Data Ware housing and why is it important for decision support?
    (c) Write the names of the data bases and information repositories on which data mining can be performed.
    (d) What do you mean by association rule mining?
    (e) Explain back propagation.
    (f) How to evaluate the accuracy of a classifier?
    (g) What can we do to secure the privacy of individuals while collecting and mining data?
    (h) Define support and confidence.
    (i) Define Lazy Learners
    (j) Explain Star schema and compare it with snow-flakes schema
    (k) Differentiate between supervised and unsupervised learning

UNIT-I

Q2   (a) Explain in detail the three data models used in Data Warehouse.   (6)
    (b) Discuss Data integration and Data transformation in detail.   (6.5)

Q3   (a) What is the difference between three main types of Data Warehouse usage: information processing, analytical processing and data mining?   (6)
    (b) Discuss the issues and benefits of data mining.   (6.5)

UNIT-II

Q4   (a) Explain the 2-tier and 3-tier data warehouse architectures in detail with suitable examples.   (6)
    (b) What is the difference between OLAP and OLTP?   (6.5)

Q5   (a) Explain single-dimensional and multi-dimensional association rules through confidence-support framework using a suitable example.   (6)
    (b) Explain Apriori algorithm of association rule mining.   (6.5)

UNIT-III

Q6   (a) What is the difference between classification and clustering? Explain the major ideas of naïve Bayesian classification.   (6)
    (b) Explain the co-relation analysis with suitable example.   (6.5)

Q7   (a) What is classification by back propagation? Discuss the steps of back propagation algorithm.   (6)
    (b) What are requirements of clustering? List major clustering methods with example. (6.5)

UNIT-IV

Q8   Discuss the applications of data mining in the following: (5x2.5=12.5)
    (a) Retail industry
    (b) Telecommunication Industry
    (c) Biological Data Analysis
    (d) Social Network analysis
    (e) Information package

Q9   (a) Write the steps to mine spatial and multimedia databases.   (6)
    (b) Explain difference between Temporal Databases and Sequential databases.   (6.5)

****

P