Code No: 07A60506 **R07** **Set No. 2**

III B.Tech II Semester Examinations,APRIL 2011
DATA WAREHOUSING AND DATA MINING
Information Technology

Time: 3 hours **Max Marks: 80**

Answer any FIVE Questions
All Questions carry equal marks

⋆ ⋆ ⋆ ⋆ ⋆

1. (a) Briefly discuss about data integration.
   (b) Briefly discuss about data transformation. [8+8]
2. Give a detail note on classification based on concepts from association rule mining. [16]
3. (a) Discuss about Concept hierarchy.
   (b) Briefly explain about - classification of database systems. [8+8]
4. (a) Explain about the graph displays of basic statistical class description.
   (b) Briefly explain about the presentation of class comparison descriptions. [8+8]
5. Discuss about primitives for specifying a data mining task. [16]
6. (a) Explain Distance-based discretization.
   (b) Give a detail note on iceberg queries . [8+8]
7. (a) Explain competitive learning and self organizing feature maps methods to clustering.
   (b) Discuss in detail BIRCH algorithm. [8+8]
8. (a) Discuss various ways to estimate the trend.
   (b) Explain construction of a multilayered web information base. [8+8]

⋆ ⋆ ⋆ ⋆ ⋆

Code No: 07A60506 **R07** **Set No. 4**

III B.Tech II Semester Examinations,APRIL 2011
DATA WAREHOUSING AND DATA MINING
Information Technology

Time: 3 hours **Max Marks: 80**

Answer any FIVE Questions
All Questions carry equal marks

⋆ ⋆ ⋆ ⋆ ⋆

1. (a) Discuss in detail DENCLUE method for custering.
   (b) What is an outlier?Explain in brief outlier analysis. [8+8]
2. (a) What is Concept description? Explain.
   (b) What are the differences between concept description in large data bases and OLAP? [8+8]
3. (a) Write an hyperlink induced topic search algorithm.
   (b) Explain latent semantic induxing technique. [8+8]
4. (a) Discuss the various measures available to judge a classifier.
   (b) Give a note on neive Bayesian classifier. [8+8]
5. (a) What does the data warehouse provide for business analyst? Explain
   (b) How do data warehousing and OLAP related to Data mining? [8+8]
6. (a) Justify the role of data cube aggregation in data reduction process with an example.
   (b) Discuss the role of Numerosity reduction in data reduction process in detail. [8+8]
7. (a) Explain how concept hierarchies are used in mining multilevel association rule?
   (b) Give the classification of association rules in detail. [8+8]
8. Explain the syntax for the following data mining primitives:
   (a) Task-relevant data
   (b) The kind of knowledge to be mined
   (c) Background knowledge
   (d) Interestingness measures. [16]

⋆ ⋆ ⋆ ⋆ ⋆

Code No: 07A60506 **R07** **Set No. 1**

III B.Tech II Semester Examinations,APRIL 2011
DATA WAREHOUSING AND DATA MINING
Information Technology

Time: 3 hours **Max Marks: 80**

Answer any FIVE Questions
All Questions carry equal marks

⋆ ⋆ ⋆ ⋆ ⋆

1. Explain any two grid based algorithms for clustering.Discuss the merits and demerits of each. [16]
2. (a) How sclable is decision tree induction?Eplain
   (b) Discuss classification based on concept from association rule mining. [8+8]
3. Explain the following terms in detail.
   (a) Concept description
   (b) Variance and Standard deviation.
   (c) Mean, median, and mode.
   (d) Quartiles, outliers, and boxplots. [16]
4. (a) Briefly explain about the forms of Data preprocessing.
   (b) Discuss issues to be considered during data integration process. [8+8]
5. (a) Discuss data transformation from time domain to frequency domain.
   (b) Explain HITS algorithm for web structure mining. [8+8]
6. (a) Briefly discuss about specifying the kind of knowledge to be mined.
   (b) Explain the syntax for specifying the kind of knowledge to be mined. [8+8]
7. What is association analysis? Discuss cluster analysis. Explain the correlation between these two types of analysis. [16]
8. (a) Draw and explain the architecture of typical data mining system.
   (b) Differentiate OLTP and OLAP. [8+8]

⋆ ⋆ ⋆ ⋆ ⋆

Code No: 07A60506 **R07** **Set No. 3**

III B.Tech II Semester Examinations,APRIL 2011
DATA WAREHOUSING AND DATA MINING
Information Technology

Time: 3 hours **Max Marks: 80**

Answer any FIVE Questions
All Questions carry equal marks

⋆ ⋆ ⋆ ⋆ ⋆

1. (a) Discuss automatic classification of web documents .
   (b) Write about basic measures of text retrieval.
   (c) Explain mining raster databases. [5+6+5]
2. Explain in detail the major steps of decision tree classification. [16]
3. (a) Explain data mining as a step in the process of knowledge discovery.
   (b) Differentiate operational database systems and data warehousing. [8+8]
4. (a) How are association rules mined from large databases?Explain.
   (b) Explain in detail constraint based association mining. [8+8]
5. Discuss about the role of data integration and transformation in data preprocessing. [16]
6. (a) Describe why is it important to have a data mining query language.
   (b) Briefly discuss about the architectures of data mining systems. [8+8]
7. (a) Give a detail note on CLIQUE algorithm.
   (b) Discuss expectation maximization algorithm for clustering. [8+8]
8. Write short notes for the following in detail:
   (a) Attribute-oriented induction.
   (b) Efficient implementation of Attribute-oriented induction. [8+8]

⋆ ⋆ ⋆ ⋆ ⋆