

Seat
No.

--	--	--	--	--	--



DII1356

Data Warehousing & Mining (New) (1290)

P. Pages : 2

Time : Three Hours

Max. Marks : 100

Instructions to Candidates :

1. Do not write anything on question paper except Seat No.
2. Answersheet should be written with blue ink only. Graph or diagram should be drawn with the same pen being used for writing paper or black HB pencil.
3. Students should note, no supplement will be provided.
4. Attempt **any two** sub questions from each question.
5. Figure to the right indicate full marks.

1. a) What is data mining ? Discuss how evolution of database technology led to data mining Also state an application where data mining is crucial to the success of business. 10
b) Briefly compare star schema and snowflake schema. Also describe which schema is more useful and state reason behind your answer. 10
c) Describe concept hierarchy. Also state why concept hierarchies are useful in data mining. 10
2. a) What is correlation analysis ? Describe its use in Data Integration. 10
b) What is the need of data reduction ? Describe attribute subset selection methods. 10
c) List and describe data mining primitives for specifying data mining task. 10
3. a) Discuss why relevance analysis is beneficial and how it can be performed and integrated in the characterization process. 10
b) Describe statistical measures to measure dispersion of data. 10
c) What is the major issues in mining class comparisons ? Describe steps of mining class comparison. 10
4. a) What is difference between naive Bayesian classifier and Bayesian Belief Networks ? Also describe components of Bayesian Belief Networks. 10

- b) Briefly outline steps of decision tree classification. 10
- c) Describe association mining based classification and rule based classification in short. 10
- 5. a) Describe K - means clustering method. Also illustrate the strength and weakness of K - means in comparison with K - medoids algorithm. 10
- b) Discuss classification of clustering methods in short. 10
- c) Write short note on following. 10
 - i) Web Usage Mining.
 - ii) Mining Multimedia databases.
