2015
B.E. (Computer Science and Engineering)
Sixth Semester
Elective – I
CS-605C: Data Mining and Analysis

Time allowed: 3 Hours   Max. Marks: 50

NOTE: Attempt _five_ questions in all, including Question No. 1 which is compulsory and selecting two questions from each Unit.

x-x-x

I. Attempt the following:-

a) What is meant by MOLAP server? State the benefits and shortcomings of this kind of server.

b) Discuss any two objective measures of pattern 'interestingness'.

c) What is the role of data transformation techniques in data mining?

d) How is the correlation coefficient calculated? What does its value signify?

e) Differentiate between hierarchical and non hierarchical clustering with the help of suitable example.   (5x2)

UNIT - I

II. What is a multidimensional data cube? What strategy can be used for materialization of the cuboids?   (10)

III. Discuss the steps followed in the data cleaning process. Are outliers and noise always undesirable in the input data?   (10)

IV. Differentiate between t-weight and d-weight measures.   (10)

UNIT - II

V. a) Describe the process of decision tree induction. Comment on the bias-variance of the DTI.   (5)

b) Distinguish between eager and lazy classifiers.   (5)

VI. a) Discuss how clustering and classification are different. Can clustering be applied to improve classification? Elaborate.   (5)

b) Write a short note on mining time series databases along with relevant applications. (5)

VII. Explain the technique used for mining multilevel association rules from transaction databases.   (10)

x-x-x