1128
M.E. (Information Technology)
Third Semester
MEIT-315: Advanced Data Mining

Time allowed: 3 Hours   Max. Marks: 50

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

x-x-x

I.   Define the following:-

    a) Why is data warehouse said to be non-volatile collection of data?

    b) Define any two techniques of data smoothing.

    c) What do you understand by market basket analysis?

    d) Why is tree pruning required in decision tree induction?

    e) What is embedded data mining?   (5x2)

UNIT – I

II.   a) List out the steps of the data warehouse design process.

    b) Differentiate between enterprise warehouse, data mart and virtual warehouse. (5,5)

III.   a) Discuss the major features and benefits of parallel RDBMS.

    b) Distinguish between roll-up and pivot OLAP operations.   (5,5)

IV.   a) Elaborate the data reduction strategies.

    b) What is concept description? Explain data generalization and summarization. (5,5)

UNIT – II

V.   a) Elaborate the concepts of the dashboards and balance scorecards.

    b) Explain the major components of BI infrastructure.   (5,5)

VI.   a) What is a data cube? Explain the significance of dimensions and facts in data cubes?

    b) Elaborate the applications of data mining in e-commerce.   (5,5)

P.T.O.

VII.   a) Discuss any two techniques for improving the efficiency of Apriori algorithm.

b) How are strong association rules generated from frequent item sets?

c) Explain Bayesian approach for classification.   (4,2,4)

x-x-x