Printed Pages: 2 **Sub Code: NMCAE13**

Paper Id: 214417 **Roll No.** | | | | | | | | | | |

MCA
(SEM IV) THEORY EXAMINATION 2017-18
DATA WAREHOUSING & MINING

Time: 3 Hours ***Total Marks: 100***

Note: Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

1. Attempt *all* questions in brief. 2 x 10 = 20

a. Define Data warehouse.
b. What is Meta Data.
c. Explain the term Crossover and Mutation in Genetic Algorithm.
d. List Various OLAP tools.
e. Why data cleaning routines are needed?
f. What is Data discretization?
g. What is Z-score normalization?
h. What do you mean by dimensionality Reduction.
i. Define Wavelet Transforms
j. Differentiate between Supervised & Unsupervised Learning.

SECTION B

2. Attempt any *three* of the following: 10 x 3 = 30

a. Explain the concept of Client/Server Computing Model. Explain the various generation of Client/ Server Model in Detail.
b. Explain the Data Warehouse Concept. Explain the two-tier data warehouse architecture and its advantages.
c. Describe ETL tools.
d. What do you mean by decision Tree? Describe ID3 algorithm for decision tree.
e. What is Web Mining? What are the classification factors of Web Mining?

SECTION C

3. Attempt any *one* part of the following: 10 x 1 = 10

(a) Draw the Data warehouse Architecture and briefly explain its various components.
(b) Why Data Warehouse is maintained separately from Database? Differentiate between OLTP & OLAP.

4. Attempt any *one* part of the following: 10 x 1 = 10

(a) Enumerate the steps involved in mapping the data warehouse to a multiprocessor architecture.
(b) Summarize the smoothing techniques followed in data cleaning process.

5. Attempt any *one* part of the following: 10 x 1 = 10

(a) Discuss, how Statistics is useful in Data Mining.

(b) Discuss various OLAP operations. Explain how query performance can be improved by cascading operations.

6. Attempt any *one* part of the following: 10 x 1 = 10

(a) What do you mean by neural network? Explain multilayer Feed -Forward neural network.

(b) Consider five points {X1, X2, X3, X4, X5} with the following coordinates as a two dimensional sample for clustering: X1=(0,2.25); X2=(0,0.25); X3=(1.25,0); X4=(4.5,0); X5=(4.5,2.5); Illustrate the K-means partitioning algorithm (clustering algorithm) using the above data set.

7. Attempt any *one* part of the following: 10 x 1 = 10

(a) What is Data Visualization? Explain various data visualization techniques. Explain Q-Q Plot.

(b) Discuss the architecture of Multimedia Data Mining.