Register Number:……..

MCA DEGREE EXAMINATIONS MAY/JUNE 2013

Fourth Semester

MASTER OF COMPUTER APPLICATIONS

MCA528: Data Warehousing and Data Mining

Time: Three Hours **Maximum Marks: 100**

Answer ALL Questions:-

PART A (10x2=20 Marks)

1. What is descriptive and predictive data mining?
2. Define data warehouse.
3. What is concept hierarchy? Give an example.
4. State the apriori property.
5. Distinguish between clustering and classification.
6. List out the major strengths of decision tree method.
7. Differentiate between agglomerative and divisive hierarchical clustering.
8. Define entropy.
9. List out any four applications of data mining.
10. What is web usage mining?

PART B (16 x 5 = 80 Marks)

11. a) (i) Describe the architecture of a data mining system (10)

(ii) Data Mining is a multi-disciplinary field. Justify. (6)

(OR)

b) (i) Justify: Data mining is a small but important step in Knowledge Discovery in Databases (KDD). (8)

(ii) Explain the patterns that can be mined using data mining. Give examples. (8)

12. a) (i) Explain the architecture of a data warehouse with a diagram. (10)

(ii) Explain about MOLAP servers in detail. (6)

(OR)

b) (i) List and discuss how a data warehouse is implemented. (10)

(ii) Explain the various OLAP operations in a multidimensional data model (6)

13. a) Explain the Apriori algorithm. Suppose you have to apply the algorithm with the assumption that the minimum support is 40% to the following set of 10 transactions that involves purchases of items A, B, C, D, E, F, and G.

T1 = {A, D, E} T2 = {A, D, F} T3 = {A, E, F} T4 = {A, B, D, F}
T5 = {B, D, F} T6 = {A, D, E, G} T7 = {A, B, D, F} T8 = {A, B, D, F, G}
T9 = {B, D, E, G} T10 = {A, D, E}

Indicate how Apriori's Large set Generation algorithm works for the example. Indicate what candidate items sets will be generated in each pass and which remains in the candidate set after pruning. Explain the process of generating association rules from the above result.

(OR)

b) What do you mean by data pre-processing? Explain the various pre-processing techniques with examples.

14. a) (i) Explain the working of a Naïve Bayesian Classifier with an example . (8)

(ii) Explain k-means cluster analysis with an example. (8)

(OR)

b) What is a decision tree? Explain the method of decision tree classification with an example.

15. a) (i) Explain how a spatial data warehouse is constructed. (8)

(ii) Discuss how data mining could be used in retail industry. (8)

(OR)

b) Explain text mining in detail giving appropriate examples.

****