APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

THIRD SEMESTER M. TECH DEGREE EXAMINATION

Computer Science & Engineering

(Computer Science & Systems Engineering)

04CS7411—Data Mining

Max. Marks: 60 Duration: 3 Hours

PART A

Answer All Questions

Each question carries 3 marks

1. Explain any 3 data mining tasks.
2. What do you mean by genetic algorithms?
3. Differentiate linear regression and non-linear regression.
4. How distance between the clusters is measured?
5. Discuss association rule, support and confidence.
6. What do you mean by divisive clustering?
7. What is web mining? Which are the three web mining tasks?
8. Explain 3 types of spatial rules.

PART B

Each question carries 6 marks

9. List and explain any 6 data mining issues.

OR

10. An airport security screening station scans the face of each passenger, and is compared to entries in a database to see if it matches with known offenders. Suggest which data mining task would help to determine it? Justify your suggestion with detailed explanation.

11. What are decision trees? Explain a simple decision tree algorithm.

OR

12. Explain an artificial neural network in detail. How feed forward neural network differs from feedback neural networks?
13. Define a 'classification problem'. Why Bayesian classification is considered as statistical based classification?

OR

14. What are support vector machines (SVM)? How SVM classifies data?
15. Define a clustering problem. Given clusters K_i and K_j, suggest some alternatives to calculate the distance between clusters.

OR

16. Illustrate hierarchical clustering algorithm with an example.
17. How quality of rules can be measured?

OR

18. A grocery store chain keeps record of items and its quantity that have sold, to find a items that are commonly purchased together. Which mining task would help to achieve this? Illustrate any standard algorithm for the same?
19. Explain the spatial data structures involved in spatial mining?

OR

20. How web content is mined using Crawlers?