



MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : PECAIML 601B Data Mining

UPID : 006920

Time Allotted : 3 Hours

Full Marks : 70

The Figures in the margin indicate full marks.

Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer any ten of the following :

[1 x 10 = 10]

- (i) What is multicollinearity?
- (ii) What is a confusion matrix? Give an example.
- (iii) Name some major sources of abundant data in various domains with examples.
- (iv) What is an attribute? What are the different types of attributes?
- (v) Define an association rule.
- (vi) What is the difference between an algorithm and a model?
- (vii) What are the different types of predictive models?
- (viii) What are the applications of KNN?
- (ix) What is LASSO regression?
- (x) How different are ROC curves for discrete and probabilistic classifier?
- (xi) Are all mined knowledge interesting?
- (xii) How are discrete and continuous attributes different?

Group-B (Short Answer Type Question)

Answer any three of the following :

[5 x 3 = 15]

2. What is frequent itemset? What is the downward closure property and upward closure property? [5]
3. What is predictive modeling? Discuss some of the most popular methods. [5]
4. What is BIC (Bayesian Information Criteria)? How to find BIC? What does a lower BIC score indicate? [5]
5. What is Bayesian belief network? Explain in details. [5]
6. What are the different steps involved in a typical web mining framework? [5]

Group-C (Long Answer Type Question)

Answer any three of the following :

[15 x 3 = 45]

7. (a) Why do we pre-process the data? [3]
- (b) Discuss briefly the various steps involved in pre-processing the data. [6]
- (c) Write in brief about data cleaning. [3]
- (d) What is imbalanced data and how to handle it? [3]
8. (a) What is prediction? Explain about linear regression method. [7]
- (b) Describe the working of SVM. [8]
9. (a) What is backward elimination? Briefly discuss with a suitable example. [5]
- (b) What is forward selection? Briefly discuss with a suitable example. [5]
- (c) What is stepwise selection? Briefly discuss with a suitable example. [5]
10. (a) What are the uses of predictive modeling? [5]
- (b) Discuss few challenges of predictive modeling. [5]
- (c) What are different types of predictive models? Describe any one. [5]
11. (a) What is ensemble learning? [3]
- (b) What is the difference between bagging and boosting? [4]
- (c) Why does bagging performs well on the low bias high variance datasets? [4]
- (d) How to calculate confusion matrix for a 2-class classification problem? [4]

*** END OF PAPER ***