

INDIAN STATISTICAL INSTITUTE

(Pattern Recognition)

MIDSEMESTER EXAMINATION

(MS-QMS)

Date: 17-02-2016
Maximum Marks: 30

Duration: 120 minutes

Note: Answer all question

1. Discuss with examples: what is data mining and what is not? [3]
2. Find and draw the BOX-and-WHISKER plot analysis for the following data set [3]
 $X=[4.3, 5.1, 3.9, 4.5, 4.4, 4.9, 5.0, 4.7, 4.1, 4.6, 4.4, 4.3, 4.8, 4.4, 4.2, 4.5, 4.4]$
3. What is variance-covariance matrix? Discuss the positive definiteness property of variance covariance matrix in the context of distance measure. [4]
4. Establish the relationship between variance-covariance matrix with correlation matrix. [2]
5. Why we go for data preprocessing? Discuss any FIVE preprocessing steps. [5]
6. Explain with example [4]
 - a. Similarity measure
 - b. Dissimilarity measure
7. Write short notes on different learning methods [4]
8. Given the following confusion matrix of a classifier, find [5]
 - a. Accuracy
 - b. Precision
 - c. Recall
 - d. Sensitivity
 - e. Selectivity of the classes

| | PREDICTED CLASS | | |
|--------------|-----------------|-----------|----------|
| | | Class=Yes | Class=No |
| ACTUAL CLASS | Class=Yes | 6954 | 46 |
| | Class=No | 412 | 2588 |