

Indian Statistical Institute, Bangalore
MS (QMS) First Year, Second Semester
Pattern Recognition

Mid-term Examination

Date: Mar 03, 2018

Maximum Marks:30

Time: 90 minutes

Note: Answer all question

1. Discuss and differentiate between the classification and clustering problems with examples. [4]
2. Differentiate between similarity and dissimilarity measures of data points with equations. [4]
3. What is generalization and over fitting in a decision making process? Differentiate them with examples. [4]
4. If X is a data set with 9 samples and 2 features each. Find the Euclidean and Mahalanobis distances between 2nd and 5th samples of X. Explain the reasons for the different distance values. [8]
$$XX = \begin{bmatrix} 4 & 5 \\ 3 & 2 \\ 3 & 2 \\ 1 & 6 \\ 0 & 2 \\ 5 & 7 \\ 7 & 9 \\ 3 & 7 \\ 2 & 1 \end{bmatrix};$$
5. Describe with examples and equations; the measures of Location, spread, Shape and dependency. [6]
6. Describe the motivations of standardization and normalization of data sets? Give one method for performing these operations. [4]