

Birla Institute of Technology & Science, Pilani
Work-Integrated Learning Programmes Division
Second Semester 2018-2019
Mid-Semester Test (EC-2 Regular)

Course No. : IS ZC464
Course Title : MACHINE LEARNING
Nature of Exam : Closed Book
Weightage : 30%
Duration : 2 Hours
Date of Exam : 10/03/2019 (FN)

No. of Pages	= 1
No. of Questions	= 4

Note:

1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

- Q.1. Answer the following questions [3 × 3 = 9]
- (a) Explain Bayes' classifier method. Give example
 - (b) Represent the classification problem as an optimization problem for training data $\{(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots, (X_N, Y_N)\}$ where (X_i, Y_i) is the i^{th} input and output pair. Give example.
 - (c) Explain the significance of parameters of slope and intercept in regression. Give example.
- Q.2. Answer the following questions [2 × 3 = 6]
- (a) What does K in K -nearest neighbor classification indicate? Explain
 - (b) Explain sensitivity analysis method of evaluating the classification algorithm.
 - (c) What do you understand by the feature vector? How is the selection of features effect classification performance. Illustrate using an example.
- Q.3. Consider a two dimensional training data belonging to two classes A and B. The feature vectors belonging to class A are $\{(5,1), (5,2), (5,4), (6,3), (6,3)\}$ and that to class B are $\{(1,0), (0,1), (1,2), (2,3), (2,0)\}$. Answer the following questions [3 x 3 = 9]
- (a) Is the data linearly separable? Explain why.
 - (b) Identify the appropriate decision boundary and write its parametric equation. What is the threshold for classification?
 - (c) Classify a test feature vector (4.5,3.5) using the above decision boundary based on the method discussed in the class. Explain all steps of mathematical computation.
- Q.4. Consider vehicle recognition problem for recognizing vehicle class of cars and buses using decision tree method. The attributes of a set of training data are shape, size and capacity of the vehicles. A total of 15 observations $\{X_1, X_2, X_3, \dots, X_{15}\}$ were taken with respect to the attribute values and their corresponding vehicle classes. The observations were collected for binary classification of vehicles into two vehicle classes - car and bus. The attribute *size* divided the training data as per the following figure. Compute the total information content of the data before splitting using size attribute. Also compute the remainder and Gain of size attribute. [6]

