

MANIPAL INSTITUTE OF TECHNOLOGY

A Constituent unit of MAHE. Manipal

# 2<sup>nd</sup> SEMESTER M.TECH. (COMPUTER SCIENCE & ENGINEERING) END SEMESTER EXAMINATIONS, April 2018

## SUBJECT: MACHINE LEARNING (PROGRAME ELECTIVE) [CSE 5255]

### REVISED CREDIT SYSTEM

### (23/04/2018)

Time: 3 Hours

MAX. MARKS: 50

2M

4M

#### Instructions to Candidates:

- ✤ Answer ALL the questions.
- Missing data may be suitably assumed.
- 1A. Define the terms 'Machine Learning', and 'Concept Learning'. Distinguish between machine learning and deep learning. Name the models for both Machine Learning and Deep Learning.
  4M
- **1B.** Write any four applications of Machine Learning. Distinguish between Fisher's Linear Discriminant and Principal Component Analysis.
- 1C. Consider a test point P which is at a distance of 1(unit) from the closest point. This sample is of Class 1. P is at a distance of 2 from the next closest point which belongs to Class 2. P is at a distance of 3 from the next closest point which is of Class 3. P is at a distance of 4 from the next closest point which is of Class 1. P is at a distance of 5 from the 5th closest point which is of Class 2. Let there be 5 classes. Let m=2. Use the Fuzzy K nearest neighbor method and find the class of P.
- 2A. Define the term discriminant function. Model the two components of probability of error in two class classification process. Show that how diagrammatically the p (error) is least when the decision boundary at a point  $g_1(x) = g_2(x)$ ? Where,  $g_1(x)$  and  $g_2(x)$  are two discriminant functions. 4M
- 2B. Write the models of the Bayes and Naïve Bayes Classifiers. 2M
- 2C. Consider the data given in the Table 1. We have a new sample money = 90, has-exams=yes, and weather=fine. Use the Naïve Bayes classifier to classify this sample as either belonging to goes-to-movie=yes or goes-to-movie=no.

Money	Has-exams	weather	Goes-to-movie
25	no	fine	no
200	no	hot	yes
100	no	rainy	no
125	yes	rainy	no
30	yes	rainy	no
300	yes	fine	yes
55	yes	hot	no
140	no	hot	no
20	yes	fine	no
175	yes	fine	yes
110	no	fine	yes

Toble 1

- **3A.** What is clustering? Why is Clustering Important? Write two applications of clustering. **3M**
- 3B. Consider the two-dimensional data set of 16 labelled samples given below. Let the two classes be labelled as "X" and "O". Let the threshold be 5 Units. X : (1, 1), (1, 2), (2, 1), (2, 2), (1, 5), (1, 6), (2, 5), (2, 6) O : (6, 1), (6, 2), (7, 1), (7, 2), (6, 6), (6, 7), (7, 6), (7, 7) Let P = (2, 3) be a test sample which needs to be classified using the nearest neighbor method of classification on the above 16 labelled samples after forming clusters. (i)Find its nearest neighbor and the right class label using nearest neighbor method of classification.(ii) Also find the centroids using clustering approach and show that clustering procedure can reduce the number of distance values to be computed from the test sample p to all the samples by using only centroids or cluster representatives. Also comment on extent of space optimization using only cluster representatives or centroids.
- **3C.** Write K-means Clustering Algorithm.
- **4A.** Define the terms Entropy and Information gain in generating decision trees? Which attribute is the best attribute?
- **4B.** Write ID3 algorithm for Decision Tree construction. Consider the following bagging problem to find the class label of the given test sample. The problem consists of the samples:s1 : (1,1.25,1), s2 : (1,1,1), s3 : (1.5,0.75,1), s4: (2,1,1), s5 : (1,3,2) s6 : (1,4,2), s7 : (1.5,3.5,2), s8 : (2,3,2), s9 : (4,2,3) s10 : (4.5,1.5,3), s11 : (5,1,3) s12 : (5,2,3). Let us have a test sample at (3, 2). Use the nearest neighbor method to find the class label of the test sample using bagging procedure.
- **4C.** Distinguish between Radial Basis Function and multilayer perceptron. Explain working of back propagation algorithm. Why it is called Backpropogation?
- 5A. Discuss support vectors, discriminant function and maximum margin of support vector machine. Write any three properties of support vector machines.
  4M
- **5B.** Discuss the working of a radial basis function network and explain how radial basis function learns boolean functions.
- **5C.** Write SOM algorithm and briefly explain the working of genetic algorithm. **3M**

2M

3M

4M

3M

3M