M.Sc. RNLKWC-/CS-401/22

2022
COMPUTER SCIENCE
M.Sc. Fourth Semester End Examination - 2022
PAPER - CS-401
Machine Learning

Full Marks : 50   *Time : 2 hours*

The figures in the right-hand margin indicate marks.
Candidates are required to give their answers in their own
words as far as practicable.
Illustrate the answers wherever necessary.

Group-A

1. **Answer any four questions :**   4×2=8

a) Explain the difference between Classification and Regression.

b) What is binomial classification and Multinomial classification.

c) Explain overfitting briefly with suitable example.

   b) Find the optimal number of cluster.

(2)

d) What do you mean by concept learning?

e) What is meant by decision boundary? What is nonlinear decision boundary?

f) Define LMS weight update rule.

Group - B

Answer any four questions :   4×4=16

2. What are the three stages to build the hypotheses or model in machine learning? What is 'Training set' and 'Test set'?
           2+2

3. What is a Neural Network? Compare between biological and artificial neuron.   1+2

4. Why does a single perceptron cannot simulate simple XOR function? Explain.
Calculate the dissimililarity between two data points $\times 1(2,3,4)$ and $\times 2(4,3,5)$
Using Euclidian distance (b) Manhattan Distance   2+2

5. Define VC dimension. Show that an axis aligned rectangle can shatter 4 points in 2 dimension.

6. Explain the steps involved in expectation maximization algorithm.   4

7. Explain the basic elements of a Hidden Mankov Mudel (HMM). List any two applications of HMM.   2+2

Group - C

Answer any two questions :

8. From bellow diagram generate statetransition table if there is a these sequence let's say "yesterday you were Happy and it was Sunny and today you are sad and it is also Sunny". then What is the probablity of this sequence?
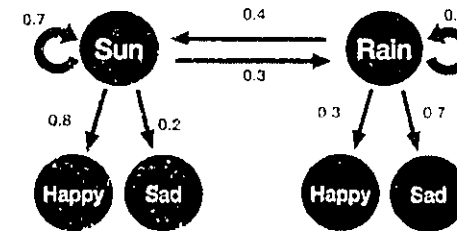
9. a) Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into clusters :

A1 (2,10), A2 (2,5), A3(8,4), A4 (5,8), A5(7,5), A6 (6,4), A7 (1,2), A8 (4,9)

10. (i) What are the benefits of pruning in decision three induction? Explain different approaches to tree pruning?

(ii) Given a set of values $x=(3,9,11,5,2)^T$ and $y=(1,8,11,4,3)^T$. Evaluate the regression coefficients

(2+3)+3

11. Explain DBSCAN algorithm for density based clustering. List out its advantages compared to K.means.

6+2