U.S.N | | | | | | | | | | |

P.E.S. College of Engineering, Mandya - 571 401

(An Autonomous Institution affiliated to VTU, Belagavi)

Seventh Semester, B.E. - Computer Science and Engineering

Semester End Examination; Dec. - 2019

Machine Learning

Time: 3 hrs *Max. Marks: 100*

***Note**: Answer **FIVE** full questions, selecting **ONE** full question from each unit.*

UNIT - I

1 a. A well poised learning problem has three major attributes associated with it- a task T a performance measure P and a training experience E. For the problems given below, mention the missing attributes as indicated in the question.

i) For a handwriting recognition learning problem, given that T-recognizing and classifying handwritten words within images E- a database of handwritten words with given classifications. What would be performance measure P? 4

ii) For a learning problem to detect fraudulent use of credit cards, given that T- recognizing fraudulent transactions identified what would be the training experience E?

b. Discuss in detail on any three issues in machine learning. 6

c. Write and explain, Find-S algorithm. Illustrate the working of FIND-S algorithm using the training examples below for the target Enjoy sport.

Ex.	Sky	Air Temp.	Humidity	Wind	Water	Forecast	Enjoy sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	change	No
4	Sunny	Warm	High	Strong	Cool	change	Yes

10

2 a. Explain the design of a learning system. 10

b. Write and explain the CANDIDATE- ELIMINATION algorithm with a suitable example. 10

UNIT - II

3 a. Consider the decision tree shown in Fig. 3(a)

i) What would be the output corresponding to the Boolean expression; (color = red^Size = small^Taste = Sweet)

ii) Write a logical expression that would correspond to the output-Apple,

iii) If the input pattern is provided as a 4-tuple always in the form {state, color, shape, size} the pattern {sweet, yellow, thin, medium} is classified as which fruit? 10

iv) Write a possible pattern in the form of 4-tuple {taste, color, shape, size} that would correspond to the output grape.

Contd…2

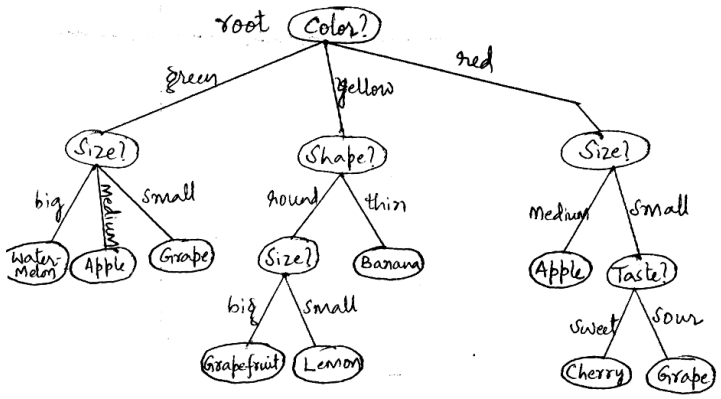

Fig 3(a)

b. Describe briefly on the issues in Decision Tree learning. 10

4 a. Consider the set of training examples below as shown in table 4(a)

i) What is the entropy of this collection of training examples with respect to the target function classification?

ii) What is the information gain of attribute A1 relative to the training samples?

iii) What is the information gain of attribute A2 relative to the training samples?

Instance	A1	A2	Classification
1	T	T	+
2	T	T	+
3	T	F	-
4	F	F	+
5	F	T	-
6	F	T	-

10

b. Write and Explain the ID3 algorithm for decision tree learning. 10

UNIT - III

5 a. Describe the characteristics of problems for which the ANN learning technique – BACKPROPOGATION is appropriate. 10

b. Derive a BACKPROPAGATION training rule for updating output unit weights. 10

6 a. With a diagram, describe briefly on perception. Illustrate how a single perception can be used to represent the Boolean functions,

i) AND

ii) OR

10

b. Explain the back propagation algorithm for feed forward network. 10

UNIT - IV

7 a. Consider the following set of training samples for the target concept Play Tennis in Table 7(a). Use Naive Bayes classifier and the training data from this table to classify and training data from this table to classify the instance:

<Outlook = Sunny, temperature = Cool, Humidity = High, Wind = Strong>

10

Contd…3

Day	Outlook	Temperature	Humidity	Wind	Play tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	No
D8	Sunny	Mild	High	Weak	No
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Hot	Normal	Weak	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

b. Describe briefly on the estimation and maximization steps of the EM algorithm. Describe how EM algorithm can be used to estimate means of K Gaussians. 10

8 a. Explain the Minimum Description Length principle. 10

b. Briefly describe on Bayes Optimal classifier.

Suppose that the Hypothesis space H has three functions h1, h2 and h3 such that P(h1/D) = 0.3, P(h2/D) = 0.3 and P(h3/D) = 0.4

i) What is the MAP hypothesis? Suppose a new instance x is encountered such that h1(x) = -1, h2(x) = -1 and h3(x) = +1. (Classified negative by h1 and h2, positive by h3) 10

ii) What are the most probable classifications of x?

iii) What is the classification generated by the MAP hypothesis?

UNIT - V

9 a. Describe K-nearest neighbor learning algorithm. What is the curse of dimensionality with reference to nearest –neighbor approach? Describe one method to overcome this. 10

b. Consider the data distribution of points in a 2-dimensional Cartesian space as shown in Table 9(b). Using K-NN learning approach (Assume K = 3) Euclidean distance as the distance metric, classify the data : (2, 4) 6

(x, y)	(2, 3)	(3, 1)	(5, 4)	(4, 3)	(1, 6)
Class	C1	C1	C2	C2	C2

Table 9(b)

c. Describe briefly on Radial Basis Function. 4

10 a. Explain the Q function. Describe an algorithm for Learning Q. 10

b. Describe any three respects in which reinforcement learning problem differs from other function approximation tasks. 6

c. Briefly describe on any strategy for improving the rate of convergence in Q learning. 4

* * *