

Time: 03 hours

[N.B. The figure in the right margin indicates the marks allocated for respective question. Split answer of a question will not be accepted.]

SECTION-A

(Answer any **three** of the followings)

1. (a) What are the purposes of studying machine learning in Computer Science and Engineering? 3
(b) Define data mining, training data, supervised and unsupervised learning. 3
(c) Write the FIND-S algorithm. Also mention its drawbacks. 4
(d) Write down the purposes of the following machine learning algorithms: 5
 - i. FIND-S
 - ii. CANDIDATE-ELIMINATION
 - iii. LIST-THEN-ELIMINATE
 - iv. ID3
2. (a) Which issues should be kept in mind to design a learning system? 3
(b) When does over-fitting occur in decision tree learning? What are the strategies to avoid over-fitting the data? 5
(c) Derive the general solution for the linear regression problem: House Price Prediction. 7
3. (a) Define Restriction biases and Preferences biases. 2
(b) A 4-input perceptron has weights 1, 2, 3, and 4. The transfer function is: if value is 79 or less call it 0; call it 1, otherwise. The inputs are 4, 10, 5 and 20 respectively. What is the classifying output? 2
(c) A hypothesis space contains three hypotheses, h_1 , h_2 and h_3 . The posteriori probabilities of these hypotheses given the training data are 0.6, 0.3 and 0.2 respectively. Suppose a new instance x is encountered, which is classified positive by h_1 , but negative by h_2 and h_3 . Using Bayes optimal classifier, find out the most probable classification of x . 4
(d) Derive the general solution to find out MAP hypotheses assuming any suitable concept learning task. Use Brute-Force MAP learning algorithm for the same. 7
4. (a) Design two distinct ANNs for the following Boolean operation with adjusted weights and coefficients: 5
 - i. AND
 - ii. OR
- (b) Consider the following *EnjoySport* artificial agent trying to infer which person will enjoy the sport based on supplied data. 10

Example	Sex	Color	Height	Nationality	EnjoySport
1	Female	Brown	Tall	Bangladesh	+
2	Male	Brown	Short	Bangladesh	-
3	Female	Black	Tall	Bangladesh	+
4	Female	Black	Tall	India	+

Your task is to build a decision tree using a suitable algorithm that predicts whether a person will enjoy the sport or not.

SECTION-B

(Answer any **three** of the followings)

1. (a) Explain the biological and artificial neural networks. 4
 - (b) Describe perceptron with an example. Which kind of problem cannot be solved by a single layer perceptron? 4
 - (c) Design a two-input perceptron that implements the Boolean function AN using gradient descent algorithm where following initial conditions are given: 7
 - i. Initial weights: $w_1=0.1$, and $w_2=0.3$
 - ii. Learning rate=0.2
 - iii. Threshold function: if the value is 0.5 or less call it 0; call it 1, otherwise.
-
2. (a) What is the role of weight and activation function in ANN? Draw the various type of activation functions used in ANN. 3
 - (b) When should you use gradient descent and when should you use normal equation to solve a linear regression problem? 3
 - (c) Define maximum-a-posteriori hypothesis and maximum likelihood hypothesis. 3
 - (d) Consider the following training data and predict the target value (yes/no) of the target concept *PlayTennis* for the new instance (Outlook=Rain, Temperature=Hot, Humidity=Normal, Wind=Strong). 6

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
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	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

3. (a) Consider a medical diagnosis problem in which there are two alternative hypotheses: (i) that the patient has a particular form of cancer, and (ii) that the patient does not. The available data is from a particular laboratory test with two possible outcomes positive and negative. You have prior knowledge that over the entire population of people only 0.008 have this disease. Furthermore, the lab test is only an imperfect indicator of the disease. The test returns a correct result in only 98% of the cases in which the disease is actually present and a correct negative result in only 97% of the cases in which the disease is not present. In other case, the test returns the opposite result. Determine the MAP hypothesis. 4
 - (b) Mention the purposes of the following machine learning algorithms: 4
 - i. Bayes optimal classifier
 - ii. GIBBS
 - iii. Naïve Bayes classifier
 - iv. Backpropagation
 - (c) Design a learning system for tic-tac-toe gaming problem. 7
-
4. (a) What are the roles of support vector machine and k-nearest neighbor in machine learning? 3
 - (b) Discuss the genetic algorithm to solve a machine learning problem. 6
 - (c) Describe briefly the face recognition procedure applying artificial neural network concept. 6