

**Seventh Semester B. E. (Computer Science and Engineering)  
Examination****Elective – II****MACHINE LEARNING**

Time : 3 Hours ]

[ Max. Marks : 60

**Instructions to Candidates :—**

- (1) All questions carry marks as indicated against them.
- (2) Assume suitable data wherever necessary.
- (3) Illustrate your answers wherever necessary with the help of neat sketches.

**1. Attempt any Two questions :—**

- (a) Calculate the size of hypothesis space in the Enjoy Sport learning task given in table 1.

Example	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.	Cloudy	Warm	High	Strong	Cool	Change	Yes

**Table 1**

How would the number of possible instances and possible hypotheses increase with the addition of attribute water Current, which can take on the values light, moderate or strong ? More generally, how does the number of possible instances and hypotheses grow with the addition of a new attribute A, which takes k possible values. 5(CO 1)

- (b) Apply candidate elimination Algorithm on training examples given in Table 1. 5(CO 1)
- (c) Find root node on decision tree using ID3 learning algorithm on Enjoy Sport Data in Table 1. 5(CO 1)

2. Attempt any **Two** questions :—

- (a) Describe Radial Basis Function (RBF) network. Solve X–OR classification problem using RBF network. 5(CO 2)
- (b) Suppose instance space  $X$  is set of Real numbers, and  $H$  is the set of intervals on the real line  $(a, b)$ . Calculate VC  $(H)$ . 5(CO 2)
- (c) Give an example in which  $k$ –Nearest Neighbors and Distance–Weighted Nearest Neighbors Algorithms, classify a query point in to two different classes. 5(CO 2)

3. Attempt any **Two** questions :—

- (a) Design a neural network and find weight matrices for X–NOR classification problem. 5(CO 2)
- (b) Consider two neurons defined by the threshold expressions  
 $w_0 + w_1 * x_1 + w_2 * x_2 > 0$ .  
Neuron A has weight values  $[1 \ 2 \ 1]$  and neuron B has weight values  $[0 \ 2 \ 1]$ . Which Neuron is more\_general\_than other neuron ? Justify your answer. 5(CO 2)
- (c) Apply Perceptron learning algorithm for 2–input NOR classification up to four iterations. Take initial weight vector  $w = [0.1 \ -0.2 \ 0.3]^T$  and  $\eta = 0.2$ . 5(CO 2)

4. Attempt any **Two** questions :—

- (a) Consider a medical diagnosis problem in which there are two alternative hypotheses :—
  - (1) that the patient, has a – particular form of cancer, and
  - (2) that the patient does not. The available data is from a particular laboratory test with two possible outcomes : positive and negative. We have prior knowledge that over the entire population of people only 0.006 have this disease.

Furthermore, the lab test is only an imperfect indicator of the disease. The test returns a correct positive 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 cases, the test returns the opposite result.

Suppose we now observe a new patient for whom the lab test returns the positive result. Should we diagnose the patient as having cancer or not ?

Suppose the doctor decides to order a second test for the same patient from the same laboratory, and suppose the second test returns a positive result as well. What are the the posterior probabilities of cancer and not cancer following these two tests ? Assume that the two tests are independent.

5(CO 3)

- (b) Apply naïve Bayes classifier to the data given in Table 2, and classify the following novel instance.

< Outlook = sunny , Temperature = cool , Humidity = high , Wind = strong >

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	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

Table 2

5(CO 3)

- (c) Suppose we have 101 coins in an urn, each with different tendency of landing heads. The first coin's tendency is 0.00, for the 2<sup>nd</sup> coin is 0.01, for the 3<sup>rd</sup> it is 0.02 — and for the last it is 1.00.

Suppose I pick a coin at random from the urn and I am about to toss it. What probability should I assign to it landing head ? 5(CO 2)

5. Attempt any **Two** questions :—

- (a) Describe Bayesian network using suitable example. Discuss the advantages of Bayesian networks. 5(CO 3)
- (b) Apply K-means clustering algorithm on following data and identify cluster for each individual (assume  $k = 2$ )

Object	X Weight index	Y pH value
A	1	1
B	2	1
C	4	3
D	5	4

**Table 3**

5(CO 3)

- (c) Use following dataset and perform hierarchical clustering using complete-linkage algorithm. Show dendrogram result.

	X	Y
1	4	4
2	8	4
3	15	8
4	24	4
5	24	12

**Table 4**

5(CO 3)

6. Attempt any **Two** questions :—

- (a) How Support Vector Machine can be used for classification of data ? 5(CO 4)
- (b) Describe ensemble learning with Bagging and Boosting. 5(CO 4)
- (c) Explain Hidden Markov Models in machine learning ; mention any one application of the same. 5(CO 4)