Course: Machine Learning   **Semester: II**
Program: M.Tech. CSE+AI
Time: 03 hrs.   **Max. Marks: 100**

SECTION A (20 Marks)

S. No.		Marks	CO
Q1	Why do we want to use weak learner when boosting?	4	CO1

Q2	Explain the following term with respect to computation learning.
	  a. Training error
	  b. True error
		4	CO1

Q3	Write short notes on reinforcement learning.	4	CO1
Q4	Cross validation can be used to select the number of iterations in boosting; this procedure may help reduce overfitting, Explain.	4	CO3
Q5	The correspondence between logistic regression and Gaussian Naive Bayes (with identity class co-variances) means that there is a one-to-one correspondence between the parameters of the two classifiers, Explain.	4	CO4

SECTION B (40 Marks)

Q6	Explain the PAC learning model.	10	CO4
Q7	What is the procedure of building decision tree using ID3 with Gain and Entropy. Illustrate with example.	10	CO3
Q8	How Naïve Bayes algorithm useful for learning and classifying text?	10	CO3
Q9	Why K-NN is called lazy learner? Differentiate between lazy and eager learner.	10	CO4

SECTION-C (40 Marks)

Q 10	i)What are the important objectives of machine learning? What are the basic design issues and approaches to machine learning? ii)Summarize K-means algorithm and group the points (1, 0, 1), (1, 1, 0), (0, 0, 1) and (1, 1, 1) using K-means algorithm.	20	CO5
Q11	The following data set can be used to learn a decision tree for predicting whether a person is happy(H) or sad(S) based on the color of their shoes, whether they wear a wig and numbers of ears they have.	20	CO2

Color	Wig	Num. Ears	Output
G	Y	2	S
G	N	2	S

G	N	2	S
B	N	2	S
B	N	2	H
R	N	2	H
R	N	2	H
R	N	2	H
R	Y	3	H

a. What attribute would decision tree algorithm choose to use for the root of tree.
b. Build a decision tree for classification of person's mood.
c. What would be the training set error for this data set? Express your answer in no of records misclassified.
d. Discuss different pruning techniques in decision tree.

OR

Write Short notes on following –
i) Kernels used in SVM
ii) Ada Boosting
iii) V C Dimension
iv) Case based Learning

Course: Machine Learning   **Semester: II**
Program: M.Tech. CSE+AI
Time: 03 hrs.   **Max. Marks: 100**

SECTION A (20 Marks)

S. No.		Marks	CO
Q1	Training error of 1NN classifier is zero, explain.	4	CO4
Q2	Why do we want to use weak learner when boosting?	4	CO3
Q3	Explain the inductive biased hypothesis space and unbiased learner.	4	CO1
Q4	Cross validation can be used to select the number of iterations in boosting; this procedure may help reduce overfitting.	4	CO3
Q5	The correspondence between logistic regression and Gaussian Naive Bayes (with identity class co-variances) means that there is a one-to-one correspondence between the parameters of the two classifiers, Explain.	4	CO3

SECTION B (40 Marks)

Q6	Explain the PAC learning model.	10	CO4
Q7	What is the procedure of building decision tree using ID3 with Gain and Entropy. Illustrate with example.	10	CO2
Q8	When does the concept of hinge loss come in picture in SVM? Explain with proper example and equations.	10	CO5
Q9	What do you mean by a well –posed learning problem? Explain the important features that are required to well –define a learning problem. Explain the inductive biased hypothesis space and unbiased learner	10	CO3

SECTION-C(40 Marks)

Q 10	i)What are the important objectives of machine learning? What are the basic design issues and approaches to machine learning? ii)Summarize K-means algorithm and group the points (1, 0, 1), (1, 1, 0), (0, 0, 1) and (1, 1, 1) using K-means algorithm.	20	CO4
Q11	The following data set can be used to learn a decision tree for predicting whether a person is happy(H) or sad(S) based on the color of their shoes, whether they wear a wig and numbers of ears they have.	20	CO2

Color	Wig	Num. Ears	Output
G	Y	2	S
G	N	2	S
G	N	2	S
B	N	2	S

B	N	2	H
R	N	2	H
R	N	2	H
R	N	2	H
R	Y	3	H

a. What attribute would decision tree algorithm choose to use for the root of tree.
b. Build a decision tree for classification of person's mood.
c. What would be the training set error for this data set? Express your answer in no of records misclassified.
d. Discuss different pruning techniques in decision tree.

OR

Write Short notes on following –
   i)   Kernels used in SVM
   ii)   Ada Boosting
   iii)   V C Dimension
   iv)   Case based Learning