## SPECIAL SUPPLE B.TECH/CSE/8ᵀᴴ SEM/CSEN 4264/2018

## MACHINE LEARNING
## (CSEN 4264)

**Time Allotted : 3 hrs**                                            **Full Marks : 70**

*Figures out of the right margin indicate full marks.*

*Candidates are required to answer Group A and*
*any 5 (five) from Group B to E, taking at least one from each group.*

*Candidates are required to give answer in their own words as far as practicable.*

## Group – A
## (Multiple Choice Type Questions)

1.  Choose the correct alternative for the following:                **10 × 1 = 10**

    (i)  What's the point of the added **bias** term in models that use linear combinations of inputs?
         (a) Due to the bias-variance tradeoff, it helps reduce overfitting
         (b) There is no point, it's just convention at this point.
         (c) In logistic/linear regression, it helps shape the curvature of the model and centers it at the origin.
         (d) Aids in a model's ability to fit the data (e.g affine transformations)

    (ii) What is **momentum** in gradient descent?
         (a) Momentum is simply the gradient multiplied by the feature vector, alluding to the physics connotation of mass times velocity.
         (b) Momentum is the projection of the gradient towards the minima multiplied by a scaling factor.
         (c) Momentum refers to the use of exponentially decaying moving averages of the past gradients to stabilize convergence.
         (d) Momentum refers to the use of exponentially decaying moving averages of learning rates in gradient descent to stabilize convergence.

    (iii) Which of the following is a predictive model?
         (a) Clustering                          (b) Regression
         (c) Summarization                        (d) Association rule.

    (iv) Multi Layer Perceptron (MLP) can be used as
         (a) a classifier                         (b) an Estimator
         (c) both (a) and (b)                     (d) none of (a) and (b).

(v) A poor binary classification model for detecting a rare cancer always predicts positive for presence of the disease. What can we infer about the model's performance?
(a) The model has high accuracy, maximum precision but low recall.
(b) The model has poor accuracy, poor precision, but maximum recall.
(c) The model has poor accuracy, maximum precision and minimum recall.
(d) The model has maximum accuracy, maximum precision but minimum recall.

(vi) Back propagation is the only learning technique that can be used to adjust weights in the neural network.
(a) True            (b) False            (c) Cannot say.

(vii) What do you mean by a hard margin?
(a) The SVM allows very low error in classification
(b) The SVM allows high amount of error in classification
(c) The SVM allows no error in classification
(d) None of the above.

(viii) The effectiveness of an SVM depends upon
(a) selection of Kernel            (b) Kernel Parameters
(c) Soft Margin Parameter C        (d) all of the above.

(ix) Let's say, a "Linear regression" model perfectly fits the training data (train error is zero). Now, Which of the following statement is true?
(a) You will always have test error zero
(b) You cannot have test error zero
(c) None of the above.

(x) Which of the following is true for neural networks?
(i) The training time depends on the size of the network.
(ii) Neural networks can be simulated on a conventional computer.
(a) All of the mentioned           (b) Only (i)
(c) Only (ii)                  (d) none of the mentioned.

## Group – B

2. (a) Explain various components of machine learning.

(b) Write the Perceptron Learning Algorithm (PLA) and briefly explain the working principle of the algorithm.

**6 + 6 = 12**

3. (a) Explain error and noise with example.

   (b) Derive the linear regression formula for single dependent variable minimizing the mean square error.

   (c) Classes attended by 10 students in machine learning and marks obtained in the examination are provided in the following table. Estimate the marks a student may obtain in the examination when she attended 20 classes, using linear regression.

| Sl No | Attendance | Marks | Sl No | Attendance | Marks |
|-------|-----------|-------|-------|-----------|-------|
| 1 | 28 | 43 | 6 | 28 | 39 |
| 2 | 27 | 39 | 7 | 26 | 36 |
| 3 | 23 | 27 | 8 | 21 | 36 |
| 4 | 27 | 36 | 9 | 22 | 31 |
| 5 | 24 | 34 | 10 | 28 | 37 |

**3 + 4 + 5 = 12**

## Group - C

4. (a) Define dichotomy, growth function and break point.

   (b) Calculate the growth function and break point for any N number of data points for
   (i) Positive rays: H consists of all hypotheses
       $h(x) = +1$ when $x \geq a$ and $h(x) = -1$ when $x < a$;
   (ii) Convex sets: H consists of all hypotheses in two dimensions
       $h : R2 \to \{-1, +1\}$ that are positive inside some convex set and negative elsewhere. Show all the steps to justify your answers.

   **6 + 6 = 12**

5. (a) Define VC dimension. What is the relation between VC dimension and break point?

   (b) Show that VC dimension for perceptron is 3 and 4 respectively in 2 and 3 dimensional data space.

   **4 + 8 = 12**

## Group – D

6. (a) Explain the single layer neural network (perceptron) for classification using the backpropagation algorithm for a data set D consisting of the training tuples and their associated target values.

(b) Explain why can't the XOR-problem be solved by a one-layer perceptron?

**8 + 4 = 12**

7. (a) Suppose that a credit card company decided to deploy a new system for assessing credit worthiness of its customers. The new system is using a feed-forward neural network with a supervised learning algorithm. Suggest the prerequisites (required information and processing) before the system can be used? Discuss issues associated with these requirements.

(b) Consider the following statement -"Classification can be considered as a special case of estimation" – Justify in favour or against of this statement.

**8 + 4 = 12**

## Group – E

8. Construct the primal problem and then derive the Lagrangian and its Dual for the optimization problem as defined by linear SVM – classification.

**12**

9. Write short notes on any 3 (three) of the followings:
   (i) Cross validation
   (ii) Mercer's condition on Kernel selection for Support Vector Machine
   (iii) Hoeffding's inequality
   (iv) Back propagation in artificial neural network
   (v) Precision and Recall.

**3 × 4 = 12**