

MACHINE LEARNING
(CSBS 3232)

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) When performing regression or classification, which of the following is the correct way to pre-process the data?
 - (a) Normalize the data → PCA → training
 - (b) PCA → normalize PCA output → training
 - (c) Normalize the data → PCA → normalize PCA output → training
 - (d) None of the above.
 - (ii) Choose the general limitations of the backpropagation rule among the following.
 - (a) Slow convergence
 - (b) Scaling
 - (c) Local minima problem
 - (d) All of the above.
 - (iii) Which of the following is/are true regarding an SVM?
 - (a) For two-dimensional data points, the separating hyperplane learnt by a linear SVM will be a straight line
 - (b) In theory, a Gaussian kernel SVM cannot model any complex separating hyperplane
 - (c) For every kernel function used in a SVM, one can obtain an equivalent closed form basis expansion
 - (d) Overfitting in an SVM is not a function of number of support vectors.
 - (iv) Among the following options identify the one which is false regarding regression.
 - (a) It is used for the prediction
 - (b) It discovers casual relationships
 - (c) It is used for interpretation
 - (d) It relates inputs to outputs.

- (v) Which of the following guidelines is applicable to initialization of the weight vector in a fully connected neural network?
- (a) Should not set it to zero since otherwise it will cause overfitting
 - (b) Should not set it to zero since otherwise (stochastic) gradient descent will explore a very small space
 - (c) Should set it to zero since otherwise it causes a bias
 - (d) Should set it to zero in order to preserve symmetry across all neurons.
- (vi) For a neural network, which one of these structural assumptions is the one that most affects the trade-off between underfitting (i.e., a high bias model) and overfitting (i.e., a high variance model)
- (a) The number of hidden nodes
 - (b) The learning rate
 - (c) The initial choice of weights
 - (d) The use of a constant-term unit input.
- (vii) The numerical output of a sigmoid node in a neural network
- (a) Is unbounded, encompassing all real numbers
 - (b) Is unbounded, encompassing all integers
 - (c) Is bounded between 0 and 1
 - (d) Is bounded between -1 and 1.
- (viii) What is unsupervised learning?
- (a) Features of group explicitly stated
 - (b) Number of groups may be known
 - (c) Neither feature nor number of groups is known
 - (d) None of the above.
- (ix) What is true about Machine Learning?
- (a) The main focus of ML is to allow computer systems to learn from experience without being explicitly programmed or human intervention
 - (b) ML is a type of artificial intelligence that extracts patterns out of raw data by using an algorithm or method
 - (c) Machine Learning (ML) is the field of computer science
 - (d) All of the above.
- (x) Which of the factors affect the performance of learner system does not include?
- (a) Representation scheme used
 - (b) Training scenario
 - (c) Type of feedback
 - (d) Good data structures.

Group- B

2. Calculate the regression coefficient and obtain the lines of regression for the following data.

[[CO2](Apply/10CQ)]

X	1	2	3	4	5	6	7
Y	9	8	10	12	11	13	14

12

3. Find the Regression coefficients for the given lines of regression $3X - 2Y = 5$ and $X - 4Y = 7$.

[[CO2](Apply/10CQ)]

12

Group - C

4. Consider the space of instances X corresponding to all points in the x, y plane. Give the VC dimension of the following hypothesis spaces:
- (i) H_r = the set of all rectangles in the x, y plane. That is, $H = \{((a < x < b) \wedge (c < y < d)) | a, b, c, d \in R\}$.
 - (ii) H_c = circles in the x, y plane. Points inside the circle are classified as positive examples.
 - (iii) H_t = triangles in the x, y plane. Points inside the triangle are classified as positive examples.
- [[CO3](Evaluate/HOCQ)]
12
5. (a) "For unbalanced dataset, accuracy is not a good metric for classification" – justify with an example. [[CO3](Evaluate /HOCQ)]
- (b) There are 550 samples of men or women in a given dataset. Out of 60% women, 75% are classified correctly, and the remaining women are classified as men by a machine learning algorithm. On the other hand, in the men category, 85% are classified as men, and the remaining are classified as women.
- i. Derive the confusion matrix.
 - ii. Calculate accuracy, precision, recall and F1 score. [[CO3/Apply(IOCQ)]
- 4 + (3 + 2 + 1 + 1 + 1) = 12**

Group - D

6. Explain the followings with respect to Back Propagation algorithm
- (i) Convergence
 - (ii) Number of Hidden Layer
 - (iii) Stopping Criterion
 - (iv) Learning rate.
- [[CO4](Understand/LOCQ)]
(4 × 3) = 12
7. (a) With suitable equations, explain any two types of activation functions used in neural networks. [[CO4](Understand/LOCQ)]
- (b) Differentiate between Gradient Descent and Stochastic Gradient Descent. [[CO2](Understand/LOCQ)]
- 6 + 6 = 12**

Group - E

8. (a) Explain how Support Vector Machine can be used for classification of linearly separable data. [[CO5, CO6](Analyze/LOCQ)]
- (b) What is the significance of optimal separating hyperplane in SVM? [[CO5](Understand/LOCQ)]
- 8 + 4 = 12**
9. (a) How can we use one-class SVM for classification? [[CO5](Understand/LOCQ)]

- (b) Prove that the kernel $K(x_1, x_2)$ is symmetric, where x_i and x_j are the feature vectors for i^{th} and j^{th} examples. [[C05](Evaluate/HOCQ)]
- (c) What are the advantages of SVM? [[C05](Understand/LOCQ)]

4 + 5 + 3 = 12

<i>Cognition Level</i>	<i>LOCQ</i>	<i>IOCQ</i>	<i>HOCQ</i>
<i>Percentage distribution</i>	<i>44.79</i>	<i>33.33</i>	<i>21.88</i>

Course Outcome (CO):

After the completion of the course students will be able to

- CSBS3232.1. Interpret the basics of the machine learning paradigm.
- CSBS3232.2. Inspect the patterns in the data using Regression techniques.
- CSBS3232.3. Contrast the complexity of Machine Learning algorithms and their limitations
- CSBS3232.4. Break down the various neural network and deep learning models.
- CSBS3232.5. Compute linear and non-linear functions using SVM
- CSBS3232.6. Apply Machine Learning algorithms in practice using real-world data

**LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.*