MACHINE LEARNING TECHNIQUES
(AEIE 4233)

Time Allotted : 2½ hrs **Full Marks : 60**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 4 (four) from Group B to E, taking one from each group.

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

Group – A

1. Answer any twelve: **12 × 1 = 12**

Choose the correct alternative for the following

(i) Which of the following is NOT a descriptive statistic?
(a) t-test (b) Mean
(c) Standard deviation (d) Range.

(ii) A normal curve is
(a) positively skewed (b) negatively skewed
(c) symmetric (d) bimodal.

(iii) If the null hypothesis is false then which of the following is accepted?
(a) Null Hypothesis (b) Positive Hypothesis
(c) Negative Hypothesis (d) Alternative Hypothesis.

(iv) The correlation for the values of two variables moving in the same direction is?
(a) Positive (b) Negative
(c) Neither positive nor negative (d) No correlation.

(v) The correlation coefficient describes
(a) Only magnitude (b) Both magnitude and direction
(c) Only direction (d) None of the preceding.

(vi) The learner is trying to predict housing prices based on the size of each house. What type of regression is this?
(a) Multivariate linear regression (b) Linear regression
(c) Logistic regression (d) Multivariate logistic regression.

(vii) What is the name of the procedure, where machine learning model parameters are changed to enhance its performance on training data set?
(a) Feature engineering (b) Model evaluation
(c) Model training (d) Model optimization.

(viii) How does SVM handle non-linearly separable data?
(a) By ignoring the data
(b) By using a linear kernel
(c) By applying the kernel trick
(d) By reducing the data size.

(ix) A perceptron is
(a) A neural network that contains feedback
(b) A single layer feed-forward neural network with pre-processing
(c) An auto-associative neural network
(d) A double layer auto-associative neural network.

(x) What is the main application of convolutional layer in a convolutional neural network (CNN)?
(a) Classification
(b) Regression
(c) Feature extraction
(d) None of the above.

Fill in the blanks with the correct word

(xi) The output of training process in machine learning is_____.

(xii) A statement made about a population for testing purpose is called _____.

(xiii) A model performs well on training data but poorly on test data due to _____.

(xiv) Machine learning is a subset of _____.

(xv) Naive Bayes Algorithm is a _____ machine learning algorithm.

Group - B

2. (a) Explain inductive learning with example. *[(CO1)(Understand/LOCQ)]*
(b) Calculate the Quartiles of the following dataset: 37, 16, 35, 13, 26, 35, 23, 26. Also determine the 60th Percentile of the above dataset? *[(CO1)(Apply/IOCQ)]*
(c) Illustrate the decision tree algorithms with an example. *[(CO1)(Understand/LOCQ)]*

2 + (3 + 2) + 5 = 12

3. (a) Explain semi-supervised learning and reinforcement learning methods. *[(CO1)(Understand/LOCQ)]*
(b) What is curse of dimensionality problem in machine learning? Explain two dimensionality reduction techniques. *[(CO1)(Understand/LOCQ)]*
(c) Describe A machine learning model is trained to predict diabetic retinopathy in patients. The test data set consists of 200 patients. The model correctly predicts positive class having retinopathy of 20 patients and negative class not having retinopathy of 120 patients. It also predicts 16 patients as negative but actually having retinopathy and 44 patients as positive but don't have retinopathy. Construct the confusion matrix based on the classifier performance and calculate precision, recall and F1 score of the model. *[(CO1)(Analyse/IOCQ)]*

3 + (1 + 2) + 6 = 12

Group - C

4. (a) State and formulate multivariate linear regression problem. Derive an expression for its solution. *[(CO2)(Analyse/IOCQ)]*

   (b) Consider a linear regression model to predict target values which suffers from over fitting problem. Suggest potential strategies to overcome this problem and improve model's predictive performance. What are the difficulties it may face? *[(CO2)(Analyse/IOCQ)]*

   (c) Explain the merits and demerits of Ridge and LASSO formulation of regression problems. *[(CO2)(Understand/LOCQ)]*

6 + 3 + 3 = 12

5. (a) Define three common performance measure of a regression model. *[(CO2)(Remember/LOCQ)]*

   (b) What are the drawbacks of linear regression model? *[(CO2)(Remember/LOCQ)]*

   (c) Distinguish between subset selection and shrinkage method. *[(CO2)(Understand/LOCQ)]*

   (d) What is principal component regression (PCR)? List the step of PCR. What are the drawbacks of PCR? *[(CO2)(Remember/LOCQ)]*

3 + 2 + 2 + (1 + 2 + 2) = 12

Group - D

6. (a) Explain the mathematical formulation of the SVM classifier. Give an outline of the method for solving the problem. *[(CO3)(Analyse/IOCQ)]*

   (b) Describe the significance of soft margin hyperplane and explain how they are computed. *[(CO3)(Understand/LOCQ)]*

   (c) Given a data set, explain the steps of performing principal component analysis (PCA) of it. *[(CO4)(Understand/LOCQ)]*

(2 + 4) + 3 + 3 = 12

7. (a) Consider a training data set consists of four training samples: (1, 4); (2, 3) in class +1 and (4, 5); (5, 6) in class -1. It is required to train a SVM classifier that learns a decision boundary leading to the largest margin from both classes. Compute the weight vector ω and bias b. *[(CO3)(Analyse/HOCQ)]*

   (b) For the above problem, find the expression of decision boundary and show it in a 2D plot along with the training samples, support vectors and margin width. *[(CO3)(Analyse/HOCQ)]*

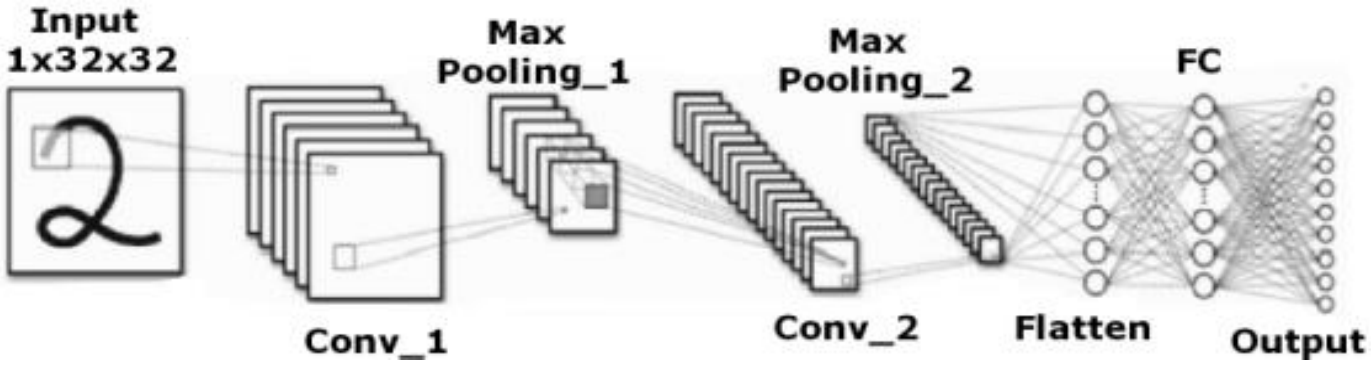
   (c) Explain how the K-means algorithm handles the assignment of data points to clusters. What criteria does it use, and how does it update the centroids? *[(CO4)(Understand/LOCQ)]*

4 + 4 + 4 = 12

Group - E

8. (a) Manish wanted to design a convolutional neural network (CNN) model the classification of digits and he considered the following network architecture as shown in figure below. He decided to use 10 and 16 filters in the first and second convolutional layers, respectively, with 5×5 filters and all operations are done

with stride 1 and no padding. All the pooling operations performed with 2×2 kernel with stride 2; and used 64 neurons in the fully connected layer (FC). Work out the output shape of each layer. *[(CO6)(Design/HOCQ)]*



(b) What will be the size of the weight matrix in between the 'flatten' and fully connected (FC) layers in the above problem? *[(CO5)(Analyse/IOCQ)]*

(c) Instead of taking 32×32 images, Manish now wants to train the same network to classify images of size 68×68. List two possible architecture changes to support this. *[(CO5)(Design/HOCQ)]*

(d) Instead of digits, Manish now wants to be able to classify handwritten alphabetic characters (26 characters). What is the minimal change in the network architecture needed in order to support this? *[(CO6)(Analyse/HOCQ)]*

6 + 2 + 2 + 2 = 12

9. (a) Suppose you have an image data set containing five different types of objects. Each of the images is of dimension 64×64×1. Design a convolutional neural network model to solve this problem. Clearly write down the shape, number of filters, kernel size, stride, pooling type and it's size in different layers. Calculate the total parameters of your designed model. *[(CO6)(Design/HOCQ)]*

(b) Suppose you want to evaluate the performance of a classifier trained to classify dogs and cats. Your test set (X_{test},Y_{test}) is such that the first m_1 images are of dogs, and the remaining images are of cats. After shuffling X_{test} and Y_{test}, you evaluate your model on it to obtain a classification accuracy a_1%. You also evaluate your model on X_{test} and Y_{test} without shuffling to obtain accuracy a_2%. What is the relationship between a_1 and a_2($<, >, \leq, \geq, =$)? Explain your answer? *[(CO5)(Analyse/IOCQ)]*

10 + 2 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	39.58	31.26	29.16