Total No. of Pages: 4 *Roll No.* 2K21/CO/259

SIXTH SEMESTER/8th B. Tech.

END TERM EXAMINATION May-2024

CO324 PATTERN RECOGNITION

Time: 3:00 Hours *Max. Marks: 50*

Note: *Attempt all questions.*
Assume suitable missing data, if any.

Q.1 [a] Consider a machine learning model trained on a dataset comprising handwritten digit images. Upon evaluation, the model exhibits a training error of 78% and a validation error of 93%. Critically analyse the implications of these results on the model's performance and propose appropriate strategies to rectify any identified shortcomings. [4] [CO1, CO2]

[b] A factory produces electronic components, of which 5% are defective and 95% are non-defective. A quality control inspector randomly selects a component and tests it. The test correctly identifies defective components 98% of the time and correctly identifies non-defective components 96% of the time. Given that the component tested is found to be defective, what is the probability that it is actually defective? [4] [CO1, CO4]

[c] Explain the problem of dimensionality and an approach to solve the problem. [2] [CO2]

Q.2 [a] Consider a dataset representing the performance of students in five subjects: Mathematics (M), Physics (P), Chemistry (C), Biology (B), and Computer Science (CS). The dataset contains the following scores for each student:

Students	M	P	C	B	CS
S1	85	90	88	92	95
S2	89	92	85	88	90
S3	82	95	91	89	92
S4	91	87	84	83	86
S5	88	86	90	85	89

Perform Principal Component Analysis (PCA) to reduce its dimensionality while retaining as much information as possible. Show the steps involved and compute the principal components along with their corresponding eigenvalues. [5] [CO3, CO4]

(b) Consider a dataset representing the heights (in inches) of students in a class. The dataset contains the following observations:

{65,68,70,72,73,66,67,69,71,68}

The heights follow a normal distribution, use maximum likelihood estimation (MLE) to estimate the mean and standard deviation of the population heights. Derive the suitable expression for MLE.

[5] [CO2, CO4]

Q.3 Attempt *any TWO* of the followings:

(a) Consider a dataset comprising 10 data points: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}. Perform Parzen window-based density estimation to estimate the probability density function (PDF) of the dataset using a Gaussian kernel with a bandwidth of h=2. Compute the PDF at x=5 using Parzen window estimation. [5] [CO3]

(b) Consider a dataset of 8 data points represented as follows:

{(1,2), (2,3), (2,5), (3,2), (4,1), (6,2), (7,3), (7,5)}

Perform Hierarchical Clustering on this dataset using the single-linkage method. Provide the dendrogram illustrating the clustering process. [5][CO3]

[c] Consider a dataset of 6 data points in two dimensions, represented as follows: {(1,2), (2,3), (3,4), (7,8), (8,9), (9,10)}

Perform Fuzzy K-means clustering on this dataset with k=2 clusters and m=2. Initialize the cluster centers {(2,3), (8,9)} and iterate until convergence. Provide the final cluster centers and the fuzzy partition matrix. [5][CO2, CO3]

Q.4 Attempt *any TWO* of the followings:

[a] Consider a simple neural network with one input layer, one hidden layer, and one output layer. The network is trained to predict the performance of students based on their study hours. The dataset contains the following data:

Study hours	2	3	4	5	6
Performance	60	75	85	90	95

Using the backpropagation algorithm with stochastic gradient descent, train the neural network to predict the performance of a student who studies for 7 hours. Assume the following initial weights and biases:

Weights: Input to Hidden Layer: 0.5; Hidden Layer to Output: 0.7

Biases: Hidden Layer: 0.3; Output Layer: 0.4

Learning Rate: 0.01 [5] [CO3, CO4]

[b] Derive the expression of gradients for backpropagation in the output layer and hidden layers. [5] [CO2, CO3]

[c] Compare and contrast different activation functions such as sigmoid, tanh, ReLU, and Leaky ReLU. [5] [CO2, CO3]

Q.5 Attempt *any four* of the followings [2.5x4=10]

[a] Explain how Simulated Annealing can improve feature selection in pattern recognition. [CO1]

[b] How does Boltzmann Learning facilitate unsupervised feature learning in pattern recognition? [CO1]

[c] How does stochastic search contribute to robustness in pattern recognition tasks? [CO2]

[d] Explain the problem of premature convergence in genetic algorithm. [CO2]

[e] Describe exploration in Particle swarm optimization with appropriate expressions. [CO2]

---Best of Luck---