Reg. No. : | | | | | | | | | | | | |

Question Paper Code: U5305

M.E. DEGREE EXAMINATION, NOV/DEC 2024

Elective

Computer Science and Engineering

21PCS505 –MACHINE LEARNIN TECHNIQUES

(Regulations 2021)

Duration: Three hours | Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 2 = 20 Marks)

1.	Describe a real-world problem that could be effectively solved using a geometric model. Explain your choice of model and its application.	CO1 - U
2.	Explain the importance of having separate training and testing datasets when developing a machine learning model.	CO1 - U
3.	Assume that you are using a distance-based method to recommend books to users based on their purchase history?	CO2- App
4.	How would you extend a binary classification model to handle multiple classes?	CO2- App
5.	Determine the optimal number of clusters, andhow would you handle features with different scales?	CO2- App
6.	Use the Singular Value Decomposition (SVD) for image compression technique	CO2- App
7.	Apply the concept of k-fold cross-validation to compute the evaluation metrics (e.g., accuracy, precision, recall).	CO2- App
8.	Describe how regularization parameters are selected using the techniques cross-validation	CO2- App
9.	How can online learning be applied to update a model with a continuous stream of data?	CO1 - U
10.	Give an example of how machine learning can be applied in healthcare for disease prediction.	CO2- App

PART B - (5 x 16 = 80 Marks)

11. (a) Develop a hidden Markov model (HMM) for speech recognition. How would you train the model and what kind of data would you use? CO2 - App (16)

Or

(b) Build a decision tree to help customers to choose a suitable car based on their preferences. What criteria would you use for splitting the nodes? CO2 -App (16)

12. (a) Design a ranking algorithm that incorporates multiple factors to rank academic papers effectively. Discuss the importance of each factor and weigh them in the ranking process. Additionally, describe the address challenges such as data availability, citation biases, and changes in research trends over time. CO2 - App (16)

Or

(b) Apply the concept of SVM to detect fraudulent transactions? Explain how you would handle class imbalance, and which kernel you might choose to improve detection accuracy. CO2 –App (16)

13. (a) An image dataset contains mixed images, where each image is a linear combination of several original images. Apply ICA to recover the original images from the mixed dataset. Explain the process and show the resulting images. CO2 - App (16)

Or

(b) Consider the following data points arranged in two concentric circles. Use Kernel K-means clustering with a suitable kernel to correctly cluster the data into two groups. Plot the clusters and explain the process. CO2 –App (16)

Point ID	X	Y
1	0.0	1.0
2	0.7	0.7
3	1.0	0.0
4	0.7	-0.7
5	0.0	-1.0
6	-0.7	-0.7
7	-1.0	0.0
8	-0.7	0.7

9	0.0	2.0
10	1.4	1.4
11	2.0	0.0
12	1.4	-1.4
13	0.0	-2.0
14	-1.4	-1.4
15	-2.0	0.0
16	-1.4	1.4

14. (a) Given a trained binary classification model, explain how you would use the ROC curve and AUC (Area Under the Curve) to evaluate its performance. You have the following predicted probabilities and true labels for 10 samples: CO2 - App (16)

Sample	Predicted Probability	True Label
1	0.8	1
2	0.6	0
3	0.7	1
4	0.3	0
5	0.9	1
6	0.4	0
7	0.2	0
8	0.5	1
9	0.1	0
10	0.75	1

Or

(b) Consider a linear regression model applied to a dataset with 100 features. Discuss the role of regularization in statistical learning theory. Specifically, compare and contrast L1 (Lasso) and L2 (Ridge) regularization in terms of their effects on model complexity, feature selection, and bias-variance tradeoff. CO2 –App (16)

15. (a) Consider a dataset with 100 labeled examples and 1000 unlabeled examples. Describe how you would train a semi-supervised learning model using self-training or co-training. CO2 - App (16)

Or

(b) Describe how you would implement a machine learning model for predictive maintenance in a manufacturing plant. Include data collection, feature engineering, model selection, and deployment steps. CO2 –App (16)