

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code: U5305

M.E. DEGREE EXAMINATION, DEC 2025

Professional Elective

Computer Science and Engineering

21PCS505 – MACHINE LEARNING TECHNIQUES

(Regulations 2021)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART - A (5 x 20 = 100 Marks)

1. (a) Analyze how error and noise affect the performance of a machine learning model in a noisy dataset. Propose strategies to mitigate these effects. CO3 Ana (20)
Or
(b) Analyze a scenario where a machine learning model exhibits high bias and low variance. What steps would you take to improve its performance? CO3 Ana (20)
2. (a) Explain the use a ranking algorithm to order product recommendations for users. What features would you consider, and how would you evaluate the effectiveness of the ranking system? CO2- App (20)
Or
(b) Given a decision tree model trained on a dataset for loan approval, Specify the extract and interpret decision rules from the tree. Provide an example of how a specific path from root to leaf translates into a decision rule. CO2- App (20)
3. (a) 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 (20)

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

Or

- (b) Given the following transaction data, use the Apriori algorithm to find all frequent item sets with a minimum support threshold of 50% and generate the association rules with a minimum confidence of 70%. CO2 -App (20)

Transaction ID	Items
1	{Bread, Milk}
2	{Bread, Diaper, Beer, Eggs}
3	{Milk, Diaper, Beer, Coke}
4	{Bread, Milk, Diaper, Beer}
5	{Bread, Milk, Diaper, Coke}

4. (a) Given a process of tuning hyper parameters for a Support Vector Machine (SVM) model. The dataset is split into a training set (70%) and a testing set (30%). Use grid search with cross-validation to find the best combination of the following hyper parameters: kernel (linear, RBF), C (0.1, 1, 10), and gamma (0.01, 0.1, 1). Explain the process and how you would evaluate the performance of the tuned model. CO2- App (20)

Or

- (b) Explain how you would use a Recurrent Neural Network (RNN) to predict the next value in a given sequence of data. Consider the following sequence and describe the process of training an RNN and making predictions

CO2 -App

(20)

Sequence Data:

[10, 20, 30, 40, 50, 60, 70, 80, 90, 100]

Detail the steps involved in preprocessing the data, defining the RNN architecture, training the model, and making predictions. Predict the next three values in the sequence and evaluate the model's accuracy.

5. (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

(20)

Or

- (b) Develop an automated customer support system using natural language processing (NLP). Apply the process to build, train, and deploy a model to understand and respond to customer inquiries.

CO2 -App

(20)

