Reg. No.																

Question Paper Code	14506

M.E. / M.Tech. - DEGREE EXAMINATIONS, APRIL / MAY 2026

Second Semester

M.E. - Big Data Analytics

24PBDPC203 - MACHINE LEARNING TECHNIQUES

Regulations -2024

Duration: 3 Hours | Max. Marks: 100

PART – A (i) (MCQ)
(16 × 1 = 16 Marks)
Answer ALL Questions

		Marks	K – Level	CO
1.	Supervised learning requires a) Labeled data b) Unlabeled data c) Random data d) No training	1	K1	CO1
2.	Overfitting occurs when a) Model fits noise b) Model ignores c) Model has low variance d) Model is simple	1	K1	CO1
3.	Expected risk in decision theory refers to a) Average loss b) Accuracy c) Entropy d) Speed	1	K1	CO1
4.	Define the function of a generative model in machine learning a) P(Y|X) b) P(X) only c) P (X, Y) d) Decision boundary only	1	K1	CO2
5.	Back propagation is used to a) Initialize weights b) Update weights c) Split data d) Reduce features	1	K1	CO2
6.	Cross entropy loss is used in a) Regression b) Classification c) Clustering d) Sampling	1	K2	CO2
7.	EM algorithm is mainly used for a) Regression b) Density estimation c) Classification trees d) Boosting	1	K2	CO3
8.	Dimensionality reduction helps in a) Increasing noise b) Reducing computational cost c) Increasing features d) Increasing variance	1	K2	CO3
9.	Infer the requirement of K-means clustering a) Number of clusters (K) predefined b) Labels c) Probabilities d) Priors	1	K2	CO3
10.	Interpret the representation of a probabilistic graphical model a) Deterministic functions b) Probabilistic dependencies among variables c) Sorting algorithms d) Linear regression	1	K2	CO4
11.	Choose the impact of conditional independence in graphical models. a) Model complexity b) Data size c) Features d) Entropy	1	K3	CO4
12.	Identify how the joint distribution is represented in Bayesian Networks. a) Sum of probabilities b) Product of conditional probabilities c) Difference of probabilities d) Entropy	1	K3	CO4
13.	Identify the role of the epsilon-greedy strategy in reinforcement learning. a) Only exploitation b) Only exploration c) Controlled exploration d) Clustering	1	K3	CO5

14.	Choose the purpose of sample complexity in machine learning. a) Training time b) Number of samples required c) Accuracy only d) Entropy	*1*	*K3*	*CO5*
15.	Identify the requirement of model-based reinforcement learning. a) Transition model b) Labels c) Clusters d) PCA	*1*	*K3*	*CO5*
16.	Choose when is the Monte Carlo reinforcement learning update the values. a) After full episode b) Immediately c) Randomly d) Never	*1*	*K3*	*CO5*

PART – A (ii) (Short Answer Questions)
(8 × 3 = 24 Marks)
Answer ALL Questions

17.	Apply probability theory to experiment its importance in solving machine learning problems.	*3*	*K3*	*CO1*
18.	Experiment with examples the distinctions between supervised, unsupervised, and reinforcement learning techniques.	*3*	*K3*	*CO1*
19.	Solve a linear regression problem using normal equation.	*3*	*K3*	*CO2*
20.	Apply cross-entropy loss to examine its role in evaluating machine learning models.	*3*	*K3*	*CO2*
21.	Utilize eigenvalues and eigenvectors to identify their significance in Principal Component Analysis (PCA).	*3*	*K3*	*CO3*
22.	Analyze the structure of a Bayesian Network.	*3*	*K4*	*CO4*
23.	Infer the concept of D-separation with an example.	*3*	*K4*	*CO4*
24.	Discover the exploration and exploitation dilemma in reinforcement learning.	*3*	*K4*	*CO5*

PART - B (Descriptive Answer Questions)
(5 × 12 = 60 Marks)
Answer ALL Questions

25.	a)	Apply Bayes Decision Theory and Maximum Likelihood Estimation to compare their performance using a suitable example.	*12*	*K3*	*CO1*
		OR			
	b)	Build a solution using decision theory concepts to address overfitting and underfitting problems with appropriate remedies.	*12*	*K3*	*CO1*
26.	a)	Construct a comparison between generative and discriminative probabilistic models based on their performance differences.	*12*	*K3*	*CO2*
		OR			
	b)	Build a comparison between neural networks and Support Vector Machines (SVM) in terms of margin, optimization, and generalization capability.	*12*	*K3*	*CO2*
27.	a)	Analyze dimensionality reduction using PCA for image compression (conceptual application).	*12*	*K4*	*CO3*
		OR			
	b)	Classify K-means and EM algorithm performance on a dataset with overlapping clusters and examine which is more suitable.	*12*	*K4*	*CO3*
28.	a)	Compare Bayesian Networks and Markov Random Fields in terms of representation, inference, and computational complexity.	*12*	*K5*	*CO4*
		OR			
	b)	Determine Hidden Markov Models and Conditional Random Fields for sequence labeling tasks and evaluate their strengths and weaknesses.	*12*	*K5*	*CO4*

29.	a)	Determine and compare Value Iteration and Policy Iteration for solving MDPs in terms of convergence and computational complexity.	*12*	*K5*	*CO5*
		OR			
	b)	Assess Monte Carlo sampling to estimate value of π using random sampling.	*12*	*K5*	*CO5*