Register No : _____

THE KAVERY ENGINEERING COLLEGE

(An Autonomous Institution)

MCA DEGREE END SEMESTER THEORY EXAMINATIONS, APR/MAY 2026

Third Semester

Computer Applications

PMCT2431 - MACHINE LEARNING

Regulations - 2024

Duration : 3 Hrs ***Maximum : 100 Marks***

PART – A(ANSWER ALL QUESTIONS) (Marks 10*2=20)

Q.No	Questions	BTL	CO	Marks
1.	What is machine learning? Mention its main types.	L1	1	2
2.	Differentiate between various types of data encountered in machine learning.	L2	1	2
3.	Describe the role of feature engineering in improving machine learning models.	L2	2	2
4.	Which of the following improves model performance? Justify your answer. a) Adding irrelevant features b) Feature transformation c) Ignoring validation d) Using less data	L2	2	2
5.	Define Bayes' rule and its significance in Bayesian learning.	L1	3	2
6.	What is a naive Bayes classifier?	L1	3	2
7.	Mention the purpose of gradient descent in model training?	L1	4	2
8.	How does dropout regularize neural networks?	L2	4	2
9.	A retail company wants to classify its customers into segments based on purchasing behavior without pre-labeled data. Which non-parametric machine learning method can they use, and why?	L2	5	2
10.	In a spam email filter using Support Vector Machines, how does the hinge loss function guide the classifier to separate spam and non-spam emails effectively?	L2	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*16 = 80)

Q.No			Questions	BTL	CO	Marks
11.	a	i	Explain the different types of human learning with examples. How can understanding these types aid in designing effective machine learning algorithms for applications such as speech recognition or autonomous vehicles?	L2	1	8
		ii	Discuss the main types of machine learning. Provide real-life examples for supervised and unsupervised learning, highlighting why the choice of learning type matters in these contexts.	L2	1	8
			Or			
	b		Explain the process of data pre-processing in machine learning. Illustrate how different types of data (e.g., categorical, numerical) are handled during pre-processing. Write an example scenario where inadequate pre-processing affected model performance and how it was corrected.	L2	1	16
12.	a		Given a dataset with imbalanced classes, demonstrate how you would select a suitable model and justify your choice based on classifier performance metrics like precision, recall, and F1-score.	L3	2	16
			Or			
	b		Consider a scenario in which you need to assess the robustness of a machine learning model in a fraud detection system. How would you evaluate the model's performance comprehensively? Discuss the use of different metrics and improving techniques in this practical context.	L3	2	16

13.	a	i	Discuss how Bayesian learning can be applied in spam filtering systems. Design a simple Bayesian classifier and describe how it adapts to new incoming email data.	L3	3	8
		ii	Given a dataset with uncertain attributes, apply Bayes' rule to update the probability of a specific event occurring. Illustrate with a step-by-step example relevant to medical diagnosis.	L3	3	8
			Or			
	b	i	How would you estimate the parameters for the maximum likelihood hypothesis in a Gaussian mixture model? Explain the steps with the formulas involved.	L3	3	8
		ii	Discuss how the EM algorithm can be applied to handle missing data in a customer segmentation problem. Describe the iterative steps involved.	L3	3	8
14.	a	i	Explain the role of a cost function in supervised learning. Using a hypothetical example, demonstrate how minimizing the cost function guides the learning process.	L3	4	8
		ii	Consider a scenario where a model exhibits high variance on validation data. Design a strategy using regularization and optimization techniques to reduce variance and improve generalization in this case study.	L3	4	8
			Or			
	b	i	Explain how a backpropagation algorithm updates weights in a multi-layer neural network. Apply this to a sample 3-layer network with given inputs and outputs.	L3	4	8
		ii	Describe various techniques to prevent overfitting in neural networks. Apply dropout regularization in a given example and explain its effect on training.	L3	4	8
15.	a	i	Analyze how the choice of k in the k-nearest neighbors algorithm affects the bias-variance tradeoff. Provide examples of how different k values impact model performance on noisy datasets.	L4	5	10
		ii	Discuss the challenges and analysis of pruning methods in reducing overfitting while maintaining model accuracy.	L4	5	6
			Or			
	b	i	A telecom company uses random forests for churn prediction. Analyze how random forests combine multiple decision trees to improve prediction accuracy and reduce overfitting. Discuss the interpretability challenges associated with this ensemble method in real-time deployment.	L4	5	10
		ii	Explain boosting in ensemble learning. Analyze how sequential training improves weak learners into a strong learner.	L4	5	6

****