Reg.No.:												

VIVEKANANDHA COLLEGE OF ENGINEERING FOR WOMEN
[AUTONOMOUS INSTITUTION AFFILIATED TO ANNA UNIVERSITY, CHENNAI]
Elayampalayam – 637 205, Tiruchengode, Namakkal Dt., Tamil Nadu.

Question Paper Code: 130003

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – NOV. / DEC. 2025
Fifth Semester
Computer Science and Technology
U23CT510 - MACHINE LEARNING
(Regulation 2023)

Time: Three Hours Maximum: 100 Marks

Answer ALL the questions

Knowledge Levels (KL)	K1 – Remembering	K3 – Applying	K5 - Evaluating
	K2 – Understanding	K4 – Analyzing	K6 - Creating

PART – A

(10 x 2 = 20 Marks)

Q. No.	Questions	Marks	KL	CO
1.	State the differences between supervised and unsupervised learning. Name at least two algorithms in each category.	2	K1	CO1
2.	Outline the concept of overfitting in a learning model.	2	K2	CO1
3.	How can linear regression be regularized to prevent overfitting? Briefly discuss different regularization techniques commonly used in linear regression.	2	K2	CO2
4.	Define Mean Absolute Error (MAE) in the context of model evaluation.	2	K2	CO2
5.	Describe Precision, Recall, F1-Score, and k-fold cross validation.	2	K1	CO3
6.	How does a Random Forest differ from a single Decision Tree?	2	K2	CO3
7.	Define hierarchical clustering. State one key distinction between hierarchical clustering and K-means clustering.	2	K1	CO4
8.	Highlight the differences between Apriori algorithm and Eclat algorithm for association rule learning.	2	K2	CO4
9.	Compare policy function and value function in reinforcement learning.	2	K2	CO5
10.	List two key differences between bagging and boosting related to how they manage error reduction.	2	K2	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.		Questions	Marks	KL	CO
11.	a)	Illustrate machine learning life cycle in detail with a neat diagram with its description. Also describe the requirement of data preprocessing and challenges with the curse of dimensionality.	13	K2	CO1
		(OR)			
	b)	Summarize the requirement of bias and variance. Clarify, how they relate to underfitting, best fitting and overfitting with the help of a learning curve. Also compare parametric and non-parametric models with suitable example.	13	K2	CO1
12.	a)	Analyze in detail the linear regression mathematically with one dependent variable and multiple independent variables. Discuss the significance of the least squares error function in the context of model estimation.	13	K3	CO2
		(OR)			
	b)	Discuss different model evaluation and performance metrics used in regression. Interpret mean absolute error (MAE), mean squared error (MSE), and R^2 Score with their mathematical formulas and explain their importance. Provide an example that demonstrates how these metrics can be used to compare regression models.	13	K3	CO2
13.	a)	Describe the working principle of Support Vector Machine (SVM) for linear classification. State the purpose of kernel function in SVM. Present one example illustrating their application.	13	K3	CO3
		(OR)			
	b)	Describe the working principle of Decision Tree algorithm for classification. What are splitting criteria, overfitting issue, and pruning techniques in decision tree? Illustrate with an example.	13	K3	CO3
14.	a)	Explain the significance of anomaly detection in unsupervised learning. Discuss any two techniques used for anomaly detection with their advantages and limitations.	13	K2	CO4
		(OR)			
	b)	Assess the impact of Principal Component Analysis (PCA) on dimensionality reduction, outlining its major strengths and key limitations in the context of data analysis and modelling. Discuss the steps involved, advantages, and limitations and also mention how PCA helps in improving machine learning model performance.	13	K2	CO4

15.	a)	Illustrate the working of reinforcement learning with suitable examples. Describe the components of an Actor-Critic Network in detail.	13	K2	CO5
		(OR)			
	b)	Outline the core concept of ensemble learning, also distinguish bagging and boosting in terms of methodology, error reduction strategy, and bias–variance management. Illustrate by citing real-world use cases for each approach	13	K2	CO5

PART – C

(1 x 15 = 15 Marks)

Q.No.		Questions	Marks	KL	CO
16.	a)	A hospital uses a diagnostic test for detecting a disease. The prevalence of the disease in the population is 1%. The test has a sensitivity of 95% (i.e., it correctly identifies 95% of diseased cases) and a specificity of 90% (i.e., it correctly identifies 90% of healthy cases). Calculate the probability using Bayes' theorem that a patient actually has the disease if the test result is positive. Discuss the role of prior probability and posterior probability in this context.	15	K3	CO3
		(OR)			
	b)	A company wants to predict the sales (in lakhs) of a product based on the advertising budget (in lakhs) spent on TV advertisements. The dataset of 5 observations is given below:	15	K3	CO2

Advertising Budget (X)	Sales (Y)
1	2
2	4
3	5
4	4
5	5

Fit a simple linear regression model of the form $Y = a + bX$. Show the step-by-step calculation of slope (b) and intercept (a). Write the final regression equation. Also discuss two real-world applications of linear regression.