

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: 13008**

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – MAY / JUNE 2024

Sixth Semester

Computer Science and Technology

U19CT612 - MACHINE LEARNING TECHNIQUES

(Regulation 2019)

Time: Three Hours

Maximum: 100 Marks

Answer ALL the questions

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

PART – A

(10 x 2 = 20 Marks)

Q.No.	Questions	Marks	KL	CO
1.	How can the Curse of dimensionality be tackled?	2	K2	CO1
2.	What is the trade off between variance and bias?	2	K3	CO1
3.	How is the Degree of the Polynomial Chosen?	2	K4	CO2
4.	Can Non-Linear Regression Handle Multiple Independent Variables?	2	K2	CO2
5.	How does Stratified k-Fold Cross-Validation Differ from k-Fold Cross-Validation?	2	K5	CO3
6.	What is Entropy in the Context of Decision Trees?	2	K3	CO3
7.	How is Support, Confidence, and Lift Defined in Association Rule Learning?	2	K1	CO4
8.	How does Imbalanced Data Impact a Confusion Matrix?	2	K4	CO4
9.	How does Negative Reinforcement Differ from Punishment?	2	K2	CO5
10.	What is the Bellman Expectation Equation for State Values?	2	K2	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.	Questions	Marks	KL	CO
11.	a) In which situations might a parametric linear Model or a non parametric linear model be preferred? (OR)	13	K2	CO1
	b) How do you Handle Missing Data in a Dataset? How do you Encode Categorical Variables? Explain.	13	K3	CO1
12.	a) Can Polynomial Regression Handle Multivariate Data? If yes then give the complete process to do so with proper examples. (OR)	13	K4	CO2
	b) What is the Importance of Initial Guesses in Non-Linear Regression? Explain	13	K4	CO2
13.	a) What is the Pitfall of Using a Single Train-Test Split for Model Evaluation? How does Cross-Validation Address the Pitfall of Overfitting in Model Evaluation? (OR)	13	K4	CO3
	b) How are Decision Trees Prone to Overfitting, and How can it be Mitigated? Explain.	13	K4	CO3
14.	a) Elaborate the Insights that can be Gained by examining Off-Diagonal Elements of a Confusion Matrix? (OR)	13	K6	CO4
	b) How are Hidden Markov Models (HMMs) used in unsupervised Learning? Illustrate a scenario in which HMM can be used.	13	K5	CO4
15.	a) Explain the Role of Escape and Avoidance in Negative Reinforcement. (OR)	13	K2	CO5
	b) Explain the Concept of Q-Values in Q-Learning. How is the Q-Value Updated in Q-Learning?	13	K4	CO5

PART – C

(1 x 15 = 15 Marks)

Q.No.	Questions	Marks	KL	CO
16. a)	<p>A manufacturing company operates a fleet of industrial machines critical to its production process. Unplanned breakdowns of these machines can result in significant downtime and financial losses. The company wants to implement a predictive maintenance system using machine learning to anticipate when a machine is likely to fail, allowing them to perform maintenance proactively.</p> <p>The company has collected historical data on machine failures, maintenance records, and sensor readings from the machines. The dataset includes information on various parameters such as temperature, pressure, vibration, and usage patterns.</p> <p>Build a machine learning model to predict machine failures and recommend maintenance actions before a breakdown occurs.</p> <ol style="list-style-type: none"> <li>How would you approach the exploration of the provided dataset?</li> <li>What preprocessing steps are necessary for preparing the data for machine learning?</li> <li>What features from the dataset do you think are most relevant for predicting machine failures?</li> <li>How would you handle missing or inconsistent data in the selected features?</li> <li>What type of machine learning model(s) would you consider for this predictive maintenance task?</li> <li>Explain the reasons behind your choice of model(s).</li> </ol> <p>(OR)</p>	2 3 2 3 2 3	K4	CO1
b)	<p>Financial institutions receive tons of requests for lending money by borrowers and making decisions for each request is a crucial task. Manually processing these requests can be a time-consuming and error-prone process, so there is an increasing demand for machine learning to improve this process by automation. The dataset contains 614 unique values for 13 columns. It contains the following columns:</p> <ul style="list-style-type: none"> <li>Loan_ID</li> <li>Gender</li> <li>Married</li> <li>Dependents</li> <li>Education</li> </ul>		K4	CO3

- Self\_Employed
  - ApplicantIncome
  - Coapplicant
  - Income
  - LoanAmount
  - Loan\_Amount\_Term
  - Credit\_History
  - Property\_Area
  - Loan\_Status
- i. How would you split the dataset into training and testing sets? 2
  - ii. What evaluation metrics would you use to assess the performance of the predictive maintenance model? 3
  - iii. Discuss the importance of hyperparameter tuning in the context of this machine learning project. 3
  - iv. How would you perform hyperparameter tuning for the chosen model? 2
  - v. Once the model is trained and validated, how would you deploy it for real-time predictions in the manufacturing environment? 2
  - vi. What considerations should be taken into account for model maintenance and updates over time? 3