	KGiSL INSTITUTE OF TECHNOLOGY (An Autonomous Institution, Affiliated to Anna University, Chennai) COIMBATORE -641 035.	Doc. Ref.	KITE/AC/IQP/76
	EXAMINATIONS - FORMS		
	END SEMESTER EXAMINATION ANSWER KEY		
	ACADEMIC YEAR: 2024 - 2025	Issue No./Date	2/15.07.2024

Reg. No.												

DEPARTMENT OF COMPUTER SCIENCE AND BUSINESS SYSTEMS

CLASS	:	II CSBS	MAX MARKS	:	60
SEMESTER	:	IV	DURATION	:	1 Hr 45 Mins
COURSE TITLE	:	MACHINE LEARNING	COURSE CODE	:	AL3451
DATE & SESSION	:	24.05.2025 & FN	EXAM	:	END SEMESTER EXAMINATION

PART – A (20 Marks)				
	ANSWER ALL QUESTIONS	CO	RBT Level	Marks
1.	**Define inductive bias in decision tree learning.** ✓ Inductive bias refers to the assumptions a learning algorithm makes to generalize beyond the training data. **(1 mark)** ✓ decision tree learning, it includes assumptions like simpler trees are preferred (Occam's Razor) and information gain is a good attribute selection measure. **(1 mark)**	CO1	U	2 (1+1)
2.	Mention different forms of learning. ✓ Supervised learning ✓ Unsupervised learning ✓ Semi-supervised learning ✓ Reinforcement learning **(2 mark)**	CO1	R	2
3.	**Name one factor that affects the convergence speed of gradient descent.** ✓ The learning rate (α) is a key factor that affects how quickly gradient descent converges. If the learning rate is too small, the algorithm converges slowly and takes many iterations. **(1 mark)** ✓ If it's too large, it may overshoot the minimum or diverge. Choosing an appropriate learning rate is crucial for effective and efficient learning. **(1 mark)**	CO2	R	2 (1+1)
4.	**If the learning rate is too large in gradient descent, what may happen?** ✓ A very large learning rate can cause the gradient descent algorithm to overshoot the optimal point during updates. **(1 mark)** ✓ This leads to fluctuations around the minimum or even complete divergence where the loss increases instead of decreasing. This instability prevents convergence and results in poor model performance. **(1 mark)**	CO2	U	2 (1+1)
5.	**List linear and nonlinear activation functions.** ✓ **Linear activation function**: f(x)=xf(x) = xf(x)=x; it's simple but cannot capture non-linearity. **(2 mark)** ✓ **Nonlinear functions:** Sigmoid: $f(x) = \frac{1}{1+e^{-x}}$ Tanh: $f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$ ReLU: $f(x) = \max(0, x)$	CO3	R	2

6.	How does CART solve the regression problems? ✓ CART (Classification and Regression Trees) handles regression by constructing a decision tree that splits data based on minimizing the mean squared error (MSE). **(1 mark)** Each internal node splits data into subsets to minimize the variance of target values in the child nodes. The final output in a leaf is the average of the target values in that node. **(1 mark)**	CO3	U	2 (1+1)
7.	**Draw the structure of an artificial single neuron based on a biological neuron.** ✓ A biological neuron has dendrites (inputs), a cell body (summation), and an axon (output). Similarly, an artificial neuron takes multiple inputs x1,x2,...,xnx_1, x_2, ..., x_nx1,x2,...,xn, multiplies them with weights w1,w2,...,wnw_1, w_2, ..., w_nw1,w2,...,wn, adds a bias, applies an activation function, and produces an output. **(2 mark)**	CO4	U	2
8.	Give the formula for accuracy. $$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Predictions}} = \frac{TP + TN}{TP + TN + FP + FN}$$ **(2 mark)**	CO4	R	2
9.	**In K-fold cross-validation, how many models are trained if K = 10?** ✓ In 10-fold cross-validation, the data is split into 10 equal parts. The model is trained 10 times, each time using 9 folds for training and 1 fold for testing. **(1 mark)** ✓ Hence, 10 models are trained, and their performance is averaged to evaluate the model more reliably. **(1 mark)**	CO5	AP	2 (1+1)
10.	**Outline the use of McNemar's test.** ✓ McNemar's test is a statistical method used to compare the performance of two classification models. It checks if the models have significantly different error rates by analyzing their disagreement on predictions. **(1 mark)** ✓ It is particularly useful when the same dataset is used for both models and helps determine if the performance difference is statistically significant. **(1 mark)**	CO5	U	2 (1+1)

PART – B (40 Marks)

		ANSWER ALL QUESTIONS	CO	RBT Level	Marks
11.	(a)	**Explain various metrics used to evaluate machine learning model performance** ✓ Accuracy: Ratio of correctly predicted instances to total instances. Good for balanced datasets. ✓ Precision: Ratio of true positives to all predicted positives. Important when false positives are costly. **(2 marks)** ✓ Recall (Sensitivity): Ratio of true positives to all actual positives. Crucial when missing a positive (false negative) is expensive. ✓ F1 Score: Harmonic mean of precision and recall. Balances the trade-off	CO1	U	16 (2+2+2+2+2+2+4)