	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