K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2025
M. Tech. Program: _Artificial Intelligence Scheme II
Regular Examination: FY Semester: II
Course Code: PCEC201 and Course Name: Deep and Reinforcement Learning
Date of Exam: 02-06-2025 Duration: 02.5 Hours Max. Marks: 60

Instructions:
(1) All questions are compulsory.
(2) Draw neat diagrams wherever applicable.
(3) Assume suitable data, if necessary.

Qu. No.	Question	Max. Marks	CO	BT level
Qu-1	Solve any **two** questions out of three: (05 marks each)	10		
a)	A robotic arm is trained to perform a pick-and-place task using DRL. Describe how you would design the state space, action space, and reward function. What challenges arise due to sparse rewards, and how can they be mitigated?	**5**	CO6	3
b)	Illustrate the agent-environment framework in Reinforcement Learning with the roles of agent, environment, states, actions, and rewards.	**5**	CO4	2
c)	How does evaluative feedback in Reinforcement Learning differ from instructive feedback in Supervised Learning? Explain with examples how these two types of feedback influence learning strategies.	**5**	CO1	2
Qu-2	Solve any **two** questions out of three: (05 marks each)	10		
a)	Explain the difference between shallow and deep neural networks. Provide an example where a shallow network might fail, but a deep one succeeds.	**5**	CO1	2
b)	Describe how dropout works as a regularization technique and its benefits in preventing overfitting	**5**	CO2	2
c)	Elaborate the DRL with one real time case study.	**5**	CO6	2
Qu-3	Solve any **two** questions out of three. (10 marks each)	20		
a)	i. Compare and contrast a CNN and a fully connected feedforward network in terms of parameter sharing and spatial hierarchy. ii. Differentiate between a Recurrent Neural Network (RNN) and a Recursive Neural Network with appropriate use cases.	**5** **5**	CO3	4 4
b)	Elaborate Temporal-Difference (TD) n reinforcement learning w.r.to: i. How does TD learning combine ideas from DP and MC methods? ii. Provide the update rule and discuss its components. iii. In what scenarios is TD learning preferred over DP and MC methods?	**10**	CO5	2
c)	Explain the back-propagation algorithm with an example. How is it used in training deep networks?	**10**	CO2	2

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2025
M. Tech. Program: _Artificial Intelligence Scheme II
Regular Examination: FY Semester: II
Course Code: PCEC201 and Course Name: Deep and Reinforcement Learning
Date of Exam: 02-06-2025 Duration: 02.5 Hours Max. Marks: 60

Qu-4	Solve any **two** questions out of three. (10 marks each)	20		
a)	A farmer went to a market and purchased a wolf, a goose, and a bag of beans. On his way home, the farmer came to the bank of a river and rented a boat. But crossing the river by boat, the farmer could carry only himself and a single one of his purchases: the wolf, the goose, or the bag of beans. If left unattended together, the wolf would eat the goose, or the goose would eat the beans. The farmer could use the boat to transfer his purchases from either bank of the river to the other. The farmer could also move across the river with the boat empty. Consider the control problem for moving all of the farmer's purchases across the river, where the current state is defined only by the following binary state variables: • which side of the river the farmer is at; • which side of the river the wolf is at; • which side of the river the goose is at; and • which side of the river the bag of beans is at. i. Define a state space for the control problem. Which is the initial state? Which states need to be defined as absorbing states? ii. Given the above information, formulate a finite Markov Decision Process (MDP) with discounting for the problem of controlling the transfer of the farmer's purchases across the river. For the transition function, give only transitions from the initial state.	**10**	CO4	2
b)	i. Compare and contrast LSTM and gated recurrent units. ii. Explain the architecture of pre trained CNN Models.	**5** **5**	CO3	4 2
c)	Elaborate Monte Carlo in the context of reinforcement learning w.r.t: i. How do MC methods differ from DP methods in terms of learning approach? ii. What is the significance of episodes in MC learning? iii. How does this algorithm estimate the value function of a policy? iv. What are the convergence properties of this algorithm?	**10**	CO5	2

****