

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

May-June 2026

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: 25-05-2026

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)	Explain the concept of Temporal-Difference (TD) Learning. How does TD Learning combine the advantages of Monte Carlo methods and Dynamic Programming?	5	CO5	U																					
b)	A hospital plans to develop an intelligent system to automatically classify medical images into multiple disease categories using Deep Learning. Design a suitable Convolutional Neural Network (CNN) model for this multiclass classification task. Explain the CNN architecture, training process, optimization techniques, regularization methods, and performance evaluation metrics used in the system.	5	CO3	Ap																					
c)	Consider the neural network with weight vector: $w=[1,-2,3,-4,2,5,-1,4,-3,2]$, Mean Squared Error (MSE) = 7 and regularization parameter $\lambda = 0.2$. Compute L1 norm, L2 norm, L1 regularized loss and L2 regularized loss.	5	CO2	Ap																					
Qu-2	Solve any TWO questions out of three: (05 marks each)	10																							
a)	Differentiate between Shallow neural network and Deep neural network.	5	CO1	U																					
b)	Consider following episode generated in Monte Carlo learning: <table border="1"><tr><td>Time step</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr><tr><td>State</td><td>S1</td><td>S2</td><td>S3</td><td>S4</td><td>S5</td><td>Terminal</td></tr><tr><td>Reward</td><td>2</td><td>4</td><td>6</td><td>8</td><td>10</td><td>0</td></tr></table> Assume the discount factor $\gamma=0.8$.Current estimate value: $V(S1)=5, V(S2)=6, V(S3)=4, V(S4)=7, V(S5)=8$ and Learning rate $\alpha=0.1$. Compute the return G_t for states S1,S2,S3,S4 and S5. Also Using the Monte Carlo learning update rule, compute the updated value estimates for all states.	Time step	1	2	3	4	5	6	State	S1	S2	S3	S4	S5	Terminal	Reward	2	4	6	8	10	0	5	CO5	Ap
Time step	1	2	3	4	5	6																			
State	S1	S2	S3	S4	S5	Terminal																			
Reward	2	4	6	8	10	0																			
c)	Compare and contrast LSTM and Gated Recurrent Units.	5	CO3	U																					

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

May-June 2026		
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: 25-05-2026	Duration: 02.5 Hours	Max. Marks: 60

Qu-3	Solve any TWO questions out of three. (10 marks each)	20		
a)	Write a short note on i) Activation functions in deep neural network ii) Applications of deep learning	5 5	CO1	U
b)	Explain any 3 optimization techniques in detail.	10	CO2	U
c)	Write a short note on i) Encoder-Decoder Sequence-to-Sequence Architectures ii) Bidirectional Recurrent Neural Network	5 5	CO3	U
Qu-4	Solve any TWO questions out of three. (10 marks each)	20		
a)	Write a short note on i) Action-Value Methods and Tracking a Non stationary Problem ii) Agent environment framework in Reinforcement learning.	5 5	CO4	U
b)	Explain Policy Evaluation and Policy Improvement with reference to Dynamic Programming.	10	CO5	U
c)	Consider the case studies of Reinforcement Learning applications and explain a suitable use case. Describe the architecture of the reinforcement learning system including agent, environment, state space, action space, reward function, and policy. Also explain the learning process and performance evaluation methods used in the system.	10	CO6	U
