

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

May-June 2026		
B. Tech. Program: Artificial Intelligence and Data Science, Scheme: II		
Regular Examination: Last Year, Semester: VIII		
Course Code: AIC801 and Course Name: Reinforcement Learning		
Date of Exam: 19/05/2026	Duration: 2.5 Hours	Max. Marks: 60

Instructions: (1) All questions are compulsory. (2) Draw neat diagrams wherever applicable. (3) Assume suitable data, if necessary.				
Q. No.	Question	Max. Marks	CO	BT level
Q 1	Solve any two questions out of three: (05 marks each)	10		
a)	State the key features of Reinforcement Learning (RL) and mention any two applications.		CO1	U
b)	What is the fundamental building block of deep neural networks, and how does it function?		CO3	U
c)	State any two advantages of policy-based methods over value-based methods.		CO4	U
Q 2	Solve any two questions out of three: (05 marks each)	10		
a)	Explain the terms "exploration" and "exploitation" in the context of Multi-arm Bandits.		CO2	U
b)	How the convolutional neural network (CNN) is used in image-related tasks.		CO3	U
c)	Explain Practical Reinforcement Learning with RL engineering and refinement process.		CO4	U
Q.3	Solve any two questions out of three. (10 marks each)	20		
a)	Define a Markov Decision Process (MDP). Explain the components of an MDP—states, actions, rewards, transition probabilities, and the discount factor.		CO2	U
b)	Name some popular deep learning frameworks and briefly explain their key features. Discuss the advantages and disadvantages of using high-level deep learning frameworks compared to low-level ones.		CO3	U
c)	Explain the concept of meta-learning algorithms adapting to different tasks or environments. Discuss the advantages and potential applications of meta-learning in reinforcement learning.		CO5	U

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

May-June 2026

B. Tech. Program: Artificial Intelligence and Data Science, Scheme: II
Regular Examination: Last Year, Semester: VIII

Course Code: AIC801 and Course Name: Reinforcement Learning

Date of Exam: 19/05/2026

Duration: 2.5 Hours

Max. Marks: 60

Q.4	Solve any two questions out of three. (10 marks each)	20		
a)	What is the standard Q-Learning algorithm, and how does it work in reinforcement learning contexts? Explain the concept of Double Q-Learning. How does it address certain issues present in standard Q-Learning?		CO2	U
b)	State the Policy Gradient Theorem and its significance in reinforcement learning. Explain how the Policy Gradient Theorem connects the gradient of the expected return with the policy parameters.		CO4	U
c)	Explain how reinforcement learning is applied in the Job-Shop Scheduling case.		CO6	U
