



# FOUNDATIONS AND APPLICATIONS OF **REINFORCEMENT LEARNING**

---

**Dr. Sudipta Majumder**

**FOUNDATIONS AND APPLICATIONS OF  
REINFORCEMENT  
LEARNING**

**Dr. Sudipta Majumder**

Senior Assistant Professor,  
Department of Computer Science and Engineering,  
Dibrugarh University Institute of Engineering and  
Technology (DUIET), Dibrugarh University.

**Kripa-Drishti Publications, Pune.**

Book Title: **Foundations and Applications of Reinforcement Learning**

Author By: **Dr. Sudipta Majumder**

Price: ₹500

1<sup>st</sup> Edition

ISBN: **978-93-48091-67-3**



Published: **July 2024**

**Publisher:**



**Kripa-Drishti Publications**

A/ 503, Poorva Height, SNO 148/1A/1/1A,  
Sus Road, Pashan- 411021, Pune, Maharashtra, India.

Mob: +91-8007068686

Email: [editor@kdpublications.in](mailto:editor@kdpublications.in)

Web: <https://www.kdpublications.in>

© Copyright **Dr. Sudipta Majumder**

All Rights Reserved. No part of this publication can be stored in any retrieval system or reproduced in any form or by any means without the prior written permission of the publisher. Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages. [The responsibility for the facts stated, conclusions reached, etc., is entirely that of the author. The publisher is not responsible for them, whatsoever.]

## **PREFACE**

In an era defined by rapid advancements in artificial intelligence, Reinforcement Learning has emerged as a cornerstone in developing intelligent systems capable of learning from interaction. Unlike traditional paradigms of supervised and unsupervised learning, reinforcement learning introduces a dynamic framework where agents adapt and evolve strategies by navigating complex environments, making decisions, and maximizing rewards over time.

This book seeks to demystify the foundational principles and cutting-edge developments in reinforcement learning. It spans from the fundamentals—introducing agents, environments, policies, and rewards—to more advanced topics like temporal difference learning, deep reinforcement learning, and practical applications in domains such as robotics, gaming, finance, and healthcare. Our focus is to equip readers with both theoretical knowledge and practical insights, fostering a comprehensive understanding of the field.

We have meticulously designed the content to cater to a diverse audience, from students and researchers exploring the basics to practitioners seeking to apply reinforcement learning to real-world challenges. Through illustrative examples, case studies, and hands-on implementations, this book bridges the gap between conceptual clarity and application-oriented expertise.

The journey of writing this book has been driven by the transformative potential of reinforcement learning. As the field continues to evolve, its implications in shaping autonomous systems, optimizing decision-making, and solving complex problems grow ever more profound. It is our hope that this book will serve as both a guide and an inspiration for those venturing into the exciting domain of reinforcement learning.

## Abbreviations

Reinforcement Learning (RL)

Temporal Difference (TD)

Intensive Care Unit (ICU)

Graphics Processing Units (GPUs)

Deep Reinforcement Learning (DRL)

Markov Decision Process (MDP)

States (S)

Actions (A)

Rewards (R)

Monte Carlo (MC)

Heating, Ventilation, and Air Conditioning (HVAC)

SARSA stands for State-Action-Reward-State-Action (SARSA)

Simulation of Urban Mobility (SUMO)

Deep Q-Learning (DQL)

Deep Q-Network (DQN)

Deep Q-Learning (DQL)

Natural Language Processing (NLP)

Convolutional Neural Network (CNN)

Proximal Policy Optimization (PPO)

A3C (Asynchronous Advantage Actor-Critic)

Deep Deterministic Policy Gradient (DDPG)

Soft Actor-Critic (SAC)

Least-Squares Temporal Difference (LSTD)

Gaussian Processes (GPs)

Support Vector Machines (SVMs)

Radial Basis Function (RBF)

Gradient Boosting Machines (GBMs)

Stochastic Gradient Descent (SGD)

# INDEX

<b>Chapter 1: Fundamentals of Reinforcement Learning.....</b>	<b>1</b>
1.1 Introduction:.....	1
1.2 Applications of Reinforcement Learning: .....	1
1.3 Challenges and Future Directions in Reinforcement Learning:.....	4
1.4 Key Concepts and Terminology:.....	7
<b>Chapter 2: Markov Decision Processes .....</b>	<b>12</b>
2.1 Introduction:.....	12
2.2 State Diagram:.....	13
2.3 Properties of Markov Chain:.....	15
2.4 Action and Reward:.....	18
2.5 Key Components of an MDP:.....	20
2.6 Goal of an MDP: .....	20
2.7 Example with Program: .....	21
2.8 Problems to Ponder: .....	29
<b>Chapter 3: Monte Carlo Methods.....</b>	<b>31</b>
3.1 Introduction:.....	31
3.2 Types of Monte Carlo Methods: .....	31
3.2.1 First-Visit Monte Carlo: .....	31
3.2.2 Every-Visit Monte Carlo: .....	32
3.3 Monte Carlo Prediction and Control: .....	33
3.3.1 Monte Carlo Prediction: .....	33
3.3.2 Monte Carlo Control: .....	34
3.4 Advantages and Disadvantages:.....	35
3.5 Applications: .....	36
3.6 Example with Program: .....	38
3.7 Case Study: Customer Purchase Behavior in a Retail Environment:.....	47
3.7.1 Pseudocode for Retail Store Monte Carlo Simulation:.....	48
3.8 Problems to Ponder: .....	50
<b>Chapter 4: Temporal Difference.....</b>	<b>53</b>
4.1 Introduction:.....	53
4.2 Application of Temporal Difference (TD) Learning:.....	54
4.3 Methods and Algorithms: .....	55
4.4 TD(0): .....	56

4.4.1 Example with Program:.....	57
4.4.2 Problems to Ponder:.....	64
4.5 TD(n): .....	66
4.5.1 Example with Program:.....	68
4.5.2 Problems to Ponder:.....	74
4.6 Q-Learning: .....	76
4.6.1 Example with Program:.....	77
4.6.2 Problems to Ponder:.....	89
4.7 SARSA:.....	90
4.7.1 Example with Program:.....	91
4.7.2 Problem to Ponder:.....	100
4.8 Comparison: .....	103
<b>Chapter 5: Deep Q-Learning .....</b>	<b>105</b>
5.1 Introduction:.....	105
5.2 How Deep Q-Learning Works?.....	106
5.3 Deep Q-Learning:.....	107
5.5 Deep Q-Network (DQN) Algorithm:.....	110
5.5 Example: .....	113
<b>Chapter 6: Function Approximation .....</b>	<b>117</b>
6.1 Introduction:.....	117
6.2 Types of Function Approximators: .....	118
6.3 Key Applications of Function Approximation:.....	120
6.4 Examples of Function Approximation in RL Algorithms: .....	121
6.4.1 Linear function Approximation: .....	123
6.4.2 Non-Linear Function Approximation: .....	126

## List of Figures

Figures Name	Page No.
Figure 1.1: Interaction between Agent and Environment	8
Figure 2.1: State diagram for a system with three states	13
Figure 2.2: State diagram for a system with three states with probabilities	14
Figure 2.3: State Diagrams with actions and rewards	14
Figure 5.1: The difference between Q-learning and deep Q-learning in evaluating the Q-value	105



## List of Tables

Table Name	Page No.
Table 2.1: Properties	18
Table 4.1: Comparison between TD(0), TD(n), Q-learning, and SARSA algorithms	103
Table 5.1: Differences Between Q-Learning, Deep Q-Learning, and Deep Q-Network:	106

## ABOUT THE AUTHOR



*Dr. Sudipta Majumder*

**Dr. Sudipta Majumder** embarked on his academic journey by completing his Bachelor of Technology (B.Tech) in Computer Science and Engineering from the North Eastern Regional Institute of Science and Technology (NERIST) in 2009. Eager to expand his knowledge, he pursued a Master of Technology (M.Tech) in Information Technology, followed by a Doctorate (Ph.D.) in Network Security, both from the prestigious North Eastern Regional Institute of Science and Technology. His doctoral research focused on advancing the field of network security, a domain he is deeply passionate about.

Dr. Majumder's primary research interests lie in the areas of internet of things, machine learning, Peer-to-Peer Networks, Wireless Networks, and Network Security. He is particularly fascinated by the challenges and opportunities presented by modern communication systems and strives to contribute innovative solutions to these fields.

Currently, Dr. Majumder serves as a Senior Assistant Professor in the Department of Computer Science and Engineering at the Dibrugarh University Institute of Engineering and Technology (DUIET), Dibrugarh University. His role allows him to guide the next generation of engineers, foster research, and collaborate with industry experts to bridge the gap between academia and real-world applications.

In addition to his academic responsibilities, Dr. Majumder is a dedicated life member of the Computer Society of India (CSI), where he actively engages with fellow professionals to promote knowledge-sharing and advancements in the field of computing.



Price: ₹ 500

**Kripa-Drishti Publications**

A-503 Poorva Heights, Pashan-Sus Road, Near Sai Chowk,  
Pune - 411021, Maharashtra, India.

Mob: +91 8007068686

Email: editor@kdpublications.in

Web: <https://www.kdpublications.in>

ISBN: 978-93-48091-67-3



9 789348 091673