1. Record Nr. UNINA9911061849303321

Autore	Nowroozi Ehsan
Titolo	Adversarial Example Detection and Mitigation Using Machine Learning / / edited by Ehsan Nowroozi, Rahim Taheri, Lucas Cordeiro
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2026
ISBN	3-031-99447-7
Edizione	[1st ed. 2026.]
Descrizione fisica	1 online resource (405 pages)
Collana	Computer Science Series
Altri autori (Persone)	Nowroozi
Disciplina	006.31
Soggetti	Artificial intelligence Machine learning Data protection - Law and legislation Cooperating objects (Computer systems) Artificial Intelligence Machine Learning Privacy Cyber-Physical Systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Preface -- Part I Foundations of Adversarial Machine Learning -- Chapter 1 A Brief Survey of Emerging Threats to AI Security -- Chapter 2 Ethical Considerations and Regulatory Standards for Adversarial Defense -- Chapter 3 Vulnerability Detection: From Formal Verification to Large Language Models and Hybrid Approaches: A Comprehensive Overview -- Part II Attacks on AI Systems -- Chapter 4 Backdoor Attacks in Text Classification: Threats, Methods, and Emerging Challenges -- Chapter 5 Biometric Template-Based Reconstruction Attack in Machine Learning -- Chapter 6 Security Weaknesses of Code Generated by Generative AI -- Chapter 7 No More Paper Tigers: A Taxonomy of Realistic Adversarial Attacks on Machine Learning based Malware Detection -- Chapter 8 Adversarial Threats to Digital Twin Technology: A Taxonomy of Vulnerabilities and Attack Surfaces -- Chapter 9 Quantum Adversarial Artificial Intelligence in Secure Internet of Things Networks -- Part III Defense Techniques and Robustness Strategies -- Chapter 10 Detecting and Mitigating Adversarial Examples

in Neural Networks: An Enhanced PGD Approach -- Chapter 11 The Role of Explainable AI (XAI) in Enhancing the Security of Machine Learning Systems Against Adversarial Attacks -- Chapter 12 Neurodevelopmental-Inspired Training Enhances Adversarial Robustness of a Primary Visual Cortex-Based Model -- Chapter 13 Evaluating and Defending Against Adversarial Attacks on LLM-Generated LSTM Models -- Chapter 14 Statistical Feature-Based Detection of Adversarial Noise and Patch Attacks in Image and Deepfake Analysis -- Chapter 15 Probabilistic Robustness in Deep Learning: A Concise yet Comprehensive Guide -- Part IV Federated Learning under Attack and Defense -- Chapter 16 Enhancing Federated Learning Security: Cluster-Based Strategies to Counter GAN-Poisoned Attacks -- Chapter 17 Defense Strategies in Federated Learning Against Adversarial Attacks -- Chapter 18 Dual Perspectives on GAN-Based Data Poisoning in Federated Learning: VagueGAN Attacks and Data Poisoning Detection -- Part V Applications and Case Studies -- Chapter 19 Cyber Risk Assessment in IT/OT Convergence using Machine Learning -- Chapter 20 Anomaly Detection Techniques in IoT Networks: Review and Comparative Analysis -- Chapter 21 Bridging the Gap from Research to Reality: Methods for Fortifying Mitigation Measures against Adversarial AI -- Index.

Sommario/riassunto

This book offers a comprehensive exploration of the emerging threats and defense strategies in adversarial machine learning and AI security. It covers a broad range of topics, from federated learning attacks, adversarial defenses, biometric vulnerabilities, and security weaknesses in generative AI to quantum threats and ethical considerations. It also brings together leading researchers to provide an in-depth and multifaceted perspective. As artificial intelligence systems become increasingly integrated into critical sectors such as healthcare, finance, transportation, and national security, understanding and mitigating adversarial risks has never been more crucial. Each chapter delivers not only a detailed analysis of current challenges, but it also includes insights into practical mitigation techniques, future trends, and real-world applications. This book is intended for researchers and graduate students working in machine learning, cybersecurity, and related disciplines. Security professionals will also find this book to be a valuable reference for understanding the latest advancements, defending against sophisticated adversarial threats, and contributing to the development of more robust, trustworthy AI systems. By bridging theoretical foundations with practical applications, this book serves as both a scholarly reference and a catalyst for innovation in the rapidly evolving field of AI security.