8th International Conference on
TEACHING, LEARNING and EDUCATION
26-28 July 2024 II Dublin, Republic of Ireland

Securing the Intelligent Future: Mitigating Attacks in AI

Nidhi Joshi Parsai[1], Rashmi Vijaywargiya[2], Sumit Jain [3]
1Assistant Professor, Dept of ISE, CMRIT, Bangalore
2Assistant Professor, Dept of CSE, IPS Indore 3Assistant Professor, Dept of CSE, SKITM Indore

Abstract

Artificial intelligence (AI) deployments become more prevalent across industries, the need for robust security measures to protect against attacks becomes paramount. This paper explores the various attack vectors that target AI systems and proposes effective strategies for mitigating these threats. We examine adversarial attacks, data poisoning, model inversion attacks, and backdoor attacks, highlighting their potential impact on critical applications such as autonomous vehicles, medical diagnosis, and financial fraud detection. To fortify AI systems, we delve into state-of-the-art defence mechanisms, including robust training techniques, anomaly detection, adversarial training, and secure model deployment practices. We also discuss the challenges and limitations associated with implementing these defences and provide insights into future research directions. By prioritizing security measures, we can ensure the responsible and secure integration of AI technologies, safeguarding the intelligent future.

Keywords: Securing, Intelligent Future Mitigating Attacks, AI Deployments