IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 7, March 2025

Adversarial Robustness of AI-Driven Claims Management Systems

Sita rRama Praveen Madugula and Nihar Malali
Independent researcher
praveenmsr@gmail.com and nihar.malali.r@gmail.com

Abstract: *Artificial intelligence (AI) has revolutionized claims management systems by streamlining processes such as fraud detection, document verification, and risk assessment, thereby enhancing operational efficiency and decision accuracy. However, AI-driven claims processing models are highly susceptible to adversarial attacks, where carefully crafted perturbations in input data can manipulate model predictions, leading to incorrect claim approvals, unjust denials, or exploitation by fraudulent actors. This study comprehensively investigates the adversarial robustness of AI-based claims management systems, analyzing different attack strategies, including evasion attacks that deceive models at inference time and poisoning attacks that corrupt training data to degrade model performance. Furthermore, it explore various defense mechanisms, such as adversarial training, robust feature extraction, uncertainty estimation, and model ensemble techniques, evaluating their effectiveness in mitigating vulnerabilities while balancing computational efficiency. Despite recent advancements, significant challenges persist in ensuring model robustness while maintaining accuracy, scalability, and compliance with evolving regulatory frameworks.*

Keywords: Claim Management Systems, Adversarial Robustness, Meta Learning, Blockchain Security

DOI: 10.48175/IJARSCT-24430