IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

Enhancing IoT Security: Implementing Deep Learning Attack Detection with Explainable AI Techniques

[1] D. Sai Venkata Abhilash, [2] Dr. N. Satheesh Kumar
[1]Research Scholar, Department of Computer Science Engineering
[2]Research Supervisor, Department of Computer Science Engineering
Sunrise University, Alwar, Rajasthan

Abstract: *The Internet of Things (IoT) connects many devices with the Internet, which facilitates life. However, this connection also increases the risk of cyber attacks. To protect these devices, we can use an attack detection system based on deep learning. Deep learning is part of artificial intelligence (AI) that helps us better understand the data. This system works by analyzing IoT device data and identifying patterns that can be harmful. When we create this deep learning system, it is essential to understand how it makes decisions. This is where it is explainable to (XAI). XAI helps us interpret the characteristics that the system uses to detect attacks. For example, you can show which data points were most important to decide if there is a threat. This is important because, in many situations, users and developers must trust that the system is making the right decisions. To implement an IoT attack detection system, we first collect devices data, such as network traffic and user behavior. Then we train the deep learning model using this data. After training, we must evaluate model performance. This implies verifying how well it detects real attacks and how many false alarms increases. Explanable techniques, such as SHAP (Shaley Adivative Explanations), can help us see what characteristics they led to certain decisions, which allows us to adjust the model for better precision.*

Keywords: Internet of Things, Artificial Inelegance (AI), Explainable AI(XAI), Cyber attacks, DNN

DOI: 10.48175/568