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 4, Issue 2, October 2024

Impact Factor: 7.53

A Novel Generative AI-Based Approach for Robust Anomaly Identification in High-Dimensional Dataset

Siddhesh Amrale
Independent Researcher
amralesiddhesh@gmail.com

Abstract: *The number of security measures put in place is wide and is in response to the rising security threats. The growing complexity of cyberattacks and the complexity of network data analysis make the conventional intrusion detection systems insufficient to locate APTs and zero-day vulnerabilities. Anomaly detection in cybersecurity has obstacles such as class imbalance, high-dimensional data, and the inability to extrapolate to shifting attack patterns. This paper showcases an AI-based approach to accurately detecting anomalies using the UNSW-NB15 benchmark dataset, which contains both typical and unusual traffic. To guarantee stable feature contribution, the suggested solution employs a thorough preprocessing pipeline, which includes data cleaning, one-hot encoding, and feature scaling. The dimensionality reduction, feature extraction for discrimination, and redundancy reduction goals are achieved through the application of PCA. In order to reduce the class imbalance, the SMOTE manifests artificial samples of minority types of attack, making model training balanced. The Generative Adversarial Network (GAN) classifier is then trained so as to differentiate the malicious and benign traffic successfully. Experimental performance is better with high precision (PRE), accuracy (ACC), recall (REC) and F1-score (F1) of 99.82, 99.75, 99.89, and 99.88, respectively compared to baseline models, which included ANN (77.51%), Decision Tree (80.5%), and KNN (97.29%). The results justify the scalability, flexibility and robustness of the proposed GAN-based framework to identify anomalies in the contemporary cybersecurity environment on time.*

Keywords*:* Cybersecurity, Anomaly detection, UNSW-NB15 dataset, Generative Artificial Intelligence, Machine learning, GAN, SMOTE.

DOI: 10.48175/IJARSCT-19900D