IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

www.ijitce.com

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 1, 2023

Ethical Hacking and Penetration Testing: Strengthening Cyber Defense with AI-Driven Breach and Attack Simulation

Charles Ubagaram,
Tata Consultancy Services
Milford, Ohio, USA
charlesubagaram17@gmail.com

Narsing Rao Dyavani,
Uber Technologies Inc, San Francisco, CA, USA
nrd3010@gmail.com

Bhagath Singh Jayaprakasam,
Cognizant Technology Solutions, Texas, USA
Bhagath.mtech903@gmail.com

Rohith Reddy Mandala,
Tekzone Systems Inc., California, USA
rohithreddymandala4@gmail.com

Venkat Garikipati,
Harvey Nash USA, Freemont, California, USA
venkat44557@gmail.com

Veerandra Kumar R,
Saveetha Engineering College, Saveetha Nagar, Thandalam, Chennai, 602105
veerandrakumar.r@panimalar.ac.in

Abstract

Ethical hacking and penetration testing are essential for identifying and mitigating security vulnerabilities in an era of increasing cyber threats. Traditional penetration testing methods, while effective, face challenges such as high costs, time-consuming assessments, and limited adaptability to evolving attack patterns. To address these limitations, this study proposes an AI-Driven Breach and Attack Simulation framework that integrates machine learning algorithms and automation techniques to simulate real-world attacks, analyze system weaknesses, and provide actionable security insights in real time. By leveraging Graph Neural Networks (GNN) for attack path prediction, BERT-based models for vulnerability detection, and Security Orchestration, Automation, and Response for risk assessment, the AI-driven Breach and Attack Simulation (BAS) system ensures continuous security evaluation while reducing false positives and response time. Experimental results demonstrate the superiority of AI-driven BAS over conventional penetration testing methods, achieving improved detection accuracy, automated remediation, and enhanced cyber resilience. This research highlights the importance of AI-powered cybersecurity solutions in proactively defending against modern cyber threats, suggesting that AI-driven BAS can revolutionize penetration testing by providing scalable, real-time, and adaptive security measures across various industries.

Keywords: Ethical Hacking, Penetration Testing, AI-Driven Breach and Attack Simulation, Cyber Defense, Vulnerability Detection, Security Automation, Machine Learning in Cybersecurity.

1. Introduction

Ethical hacking and penetration testing are essential cybersecurity practices that help organizations identify and mitigate vulnerabilities before malicious hackers exploit them [1]. These proactive security measures simulate real-world cyberattacks to assess system defenses and improve security postures [2]. Ethical hackers use various tools and techniques to uncover weaknesses in networks, applications, and hardware [3]. Companies across industries rely on penetration testing to safeguard sensitive data, prevent financial losses, and ensure compliance with cybersecurity regulations [4]. With the rise in cyber threats, traditional security measures alone are insufficient, necessitating continuous testing and improvement [5]. Ethical hacking is conducted within legal boundaries, following strict guidelines to ensure responsible vulnerability disclosure [6]. It plays a critical role in risk management by providing insights into potential attack vectors [7]. The evolving threat landscape demands advanced penetration testing methodologies to stay ahead of cybercriminals [8].

Cybersecurity vulnerabilities arise due to poor security configurations, weak passwords, and outdated software [9]. Human errors, such as accidental data exposure and phishing attacks, are among the leading causes of breaches [10]. Insufficient network monitoring and lack of encryption mechanisms leave systems exposed to cyber threats [11]. Increasing reliance on cloud computing and IoT devices introduces new security challenges [12]. Zero-day vulnerabilities in software and applications provide attackers with opportunities to exploit unknown weaknesses [13]. Organizations often fail to conduct regular security assessments, leading to unpatched vulnerabilities [14]. Social engineering tactics manipulate employees into revealing sensitive information, making businesses susceptible to breaches [15]. Cybercriminals continuously develop sophisticated attack methods, making traditional security defenses inadequate [16].

Traditional penetration testing involves manual and automated techniques to assess system security, but it is often time-consuming and costly [17]. Black-box testing simulates real-world attacks but lacks in-depth insights into internal system vulnerabilities [18]. White-box testing provides access to source code and system architecture but requires extensive expertise and resources [19]. Vulnerability scanners help detect known weaknesses but generate false positives, making it difficult to prioritize threats [20]. IDS monitor network traffic but may fail to detect APTs [21]. Security audits and compliance checks provide regulatory adherence but do not guarantee real-time protection against evolving attacks [22]. Red teaming exercises enhance security resilience but require skilled professionals and significant investment [23]. Traditional methods struggle to adapt to the increasing complexity of cyber threats and fail to provide continuous security assessment [24]. The growing demand for real-time threat detection and response necessitates a more dynamic and AI-driven approach [25].

To overcome these limitations, AI-driven BAS offers an automated, intelligent approach to cybersecurity testing. This method continuously simulates cyberattacks using machine learning algorithms to identify vulnerabilities and predict potential threats. AI-driven BAS adapts to new attack patterns in real-time, providing continuous security assessments. Unlike traditional penetration testing, it reduces manual efforts and minimizes false positives, improving accuracy. By leveraging AI, organizations can automate security testing and enhance threat intelligence. AI-driven BAS enables proactive defense mechanisms by analyzing attack paths and suggesting immediate remediation strategies. This approach ensures real-time monitoring and adaptive security measures, making cybersecurity testing more efficient. Implementing AI-driven BAS strengthens cyber defense, reduces response time, and enhances overall security resilience.

In Section 2, cybersecurity vulnerabilities and the limitations and Section 3, existing methods like IDS. Section 4 illustrates the AI-Driven Cybersecurity Framework for Threat Detection and Attack Simulation. In Section 5, Results show improved accuracy and threat mitigation, while Section 6, conclusion and future works.

2. Literature Review

Dondapati Utilized a [26] AI-driven methods such as deep learning, image analysis, and genetic data interpretation are increasingly used for lung cancer diagnosis and treatment prediction. However, traditional techniques face challenges like limited accuracy, high false positives, and computational complexity, necessitating optimized AI models for improved patient outcomes. Gudivaka proposed [27] Big data-driven methods using Hadoop-based big data analytics and machine learning improve silicon prediction in blast furnace smelting but face challenges like data integration, high costs, and real-time processing limitations.

Alagarsundaram & Carolina stated, [28] AES is implemented in cloud computing using symmetric key encryption, involving key expansion, substitution-permutation networks, and multiple transformation rounds. However, challenges include performance overhead, key management complexity, and compatibility issues across diverse cloud environments. Yallamelli [29] Integrated big data, hash graph, and cloud computing using the Kinetic methodology enhances data processing, security, and decision-making. Cloud computing offers scalability, big data enables insights, and hash graph ensures secure consensus. However, challenges like interoperability, scalability, and regulatory compliance remain.

Allur [30] studied enhances software testing using Genetic Algorithms, Hybrid Optimization (GA+PSO+ACO), and Co-evolutionary Techniques for better test data generation and path coverage. While improving efficiency and scalability, challenges include high computational cost, parameter tuning, and resource consumption. Deevi [31] studies uses WBANs, multi-biometric key generation, and DWT for secure mobile healthcare. Limitations include high computational complexity and real-time processing challenges.

3. Problem Statement

Traditional diagnostic and predictive methods face limitations in accuracy, high false positives, and computational inefficiencies [32]. The need for AI-driven solutions, optimized encryption techniques, and big data analytics is crucial to overcoming these challenges in various domains, including healthcare, manufacturing, and cloud security [33].

Furthermore, existing approaches in software testing and mobile healthcare security struggle with high computational costs, real-time processing limitations, and interoperability concerns [34]. Developing scalable, efficient, and secure methodologies is essential for improving decision-making, system reliability, and data protection.

4. AI-Driven Cybersecurity Framework for Threat Detection and Attack Simulation

The diagram represents an AI-driven cybersecurity framework for breach and attack simulation. It begins with data collection, gathering security-related information. The data then undergoes preprocessing using normalization to ensure consistency and accuracy. Next, attack path prediction using GNN helps identify potential attack routes. This is followed by vulnerability detection using BERT, which analyzes system weaknesses. The detected vulnerabilities are integrated into threat intelligence using SOAR (Security Orchestration, Automation, and Response) to automate response strategies. Finally, performance evaluation assesses the framework's effectiveness in mitigating threats, ensuring a robust and proactive security system is shown in Figure (1),

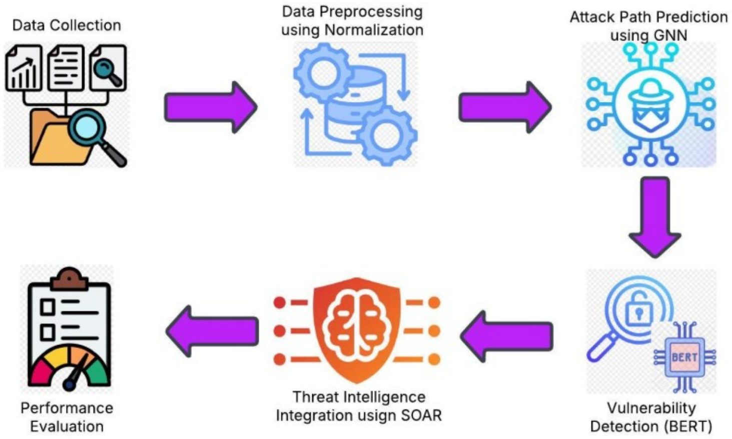
Data Collection
Data Preprocessing using Normalization
Attack Path Prediction using GNN
Vulnerability Detection (BERT)
BERT
Threat Intelligence Integration usign SOAR
Performance Evaluation

Figure 1: Intelligent Threat Detection and Attack Path Prediction Using AI

4.1 Data Collection

The Malware Detection in Network Traffic Data dataset from Stratosphere Laboratory provides labeled IoT network traffic for identifying malicious activities. It includes labels such as Attack, Benign, C&C, DDoS, File Download, and Heart Beat. It also identifies botnets like Mirai, Okiru, and Torii, along with Part of a Horizontal Port Scan for reconnaissance. This dataset is essential for AI-driven malware detection and cybersecurity research.

Data Set Link: https://www.kaggle.com/datasets/agungpambudi/network-malware-detection-connection-analysis

4.2 Data Preprocessing using Normalization

To preprocess the data, Robust Scaling and Normalization are used to handle outliers and ensure consistency. Min-Max Scaling normalizes the data within a fixed range [0,1] using the formula is indicated as Eq. (1),

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

Where, X is the original data, $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the dataset, and X′ is the scaled data. Robust Scaling, on the other hand, is based on the median and IQR to minimize the effect of outliers. It is calculated as given by Eq. (2),

$$X' = \frac{X - \text{median}(X)}{\text{IQR}(X)} \tag{2}$$

Where, the IQR helps in scaling without being affected by extreme values. These techniques improve data quality and stability for AI models in cybersecurity applications.

4.3 Attack Path Prediction using GNN

In Attack Path Prediction using GNNs, the network is represented as a graph G = (V, E) where, V consists of devices, servers, and users (nodes), and E represents connections between them (edges). GNNs learn attack paths by analyzing these connections using an adjacency matrix A that captures relationships between nodes. The node features evolve across layers using the equation is described as Eq. (3),

$$H^{(l+1)} = \sigma\left(W^{(l)}H^{(l)}A + B\right) \tag{3}$$

Where, $H^{(l)}$ is the feature matrix at layer l, $W^{(l)}$ is a trainable weight matrix, B is a bias term, and σ is an activation function (e.g., ReLU). This helps in predicting how attackers can move laterally across the system, identifying weak points for security reinforcement.

4.4 Vulnerability BERT- Based Model for Detection

In VulBERTa, the BERT model processes security logs and system configurations by tokenizing and embedding the input data as indicated as Eq. (4),

$$E = BERT(T) \tag{4}$$

Where, T represents the input text, and E is the contextual embedding capturing vulnerability-related information. The model then classifies vulnerabilities using a Softmax function, which assigns probabilities to different vulnerability types is given below in Eq. (5),

$$P(y \mid X) = \text{Softmax}(W \cdot E + b) \tag{5}$$

Where, W is the weight matrix, b is the bias term, and Softmax normalizes the scores into probabilities. This enables the model to detect vulnerabilities accurately by understanding contextual patterns in security data.

4.5 Threat Intelligence Integration using SOAR

In Threat Intelligence Integration using SOAR, threat intelligence is gathered from sources like MITRE ATT&CK, CVEs, and IDS logs to automate responses based on detected vulnerabilities. SOAR dynamically updates a risk score R to prioritize threats using the formula is indicated in Eq. (6),

$$R = \sum_{i=1}^{N} w_i \cdot S_i \tag{6}$$

Where, S_i represents the severity of each detected vulnerability, and w_i is a weight factor based on exploitability and impact. Higher risk scores indicate critical vulnerabilities requiring immediate action. By automating responses, SOAR enhances cybersecurity by reducing manual effort and improving threat mitigation efficiency.

5. Results and Discussion

This section analyzes the performance of the BERT-based model for vulnerability detection in cybersecurity, demonstrating high accuracy, precision, recall, F1-score, and ROC-AUC, ensuring reliable threat identification. Visual representations highlight its effectiveness in detecting various cyber threats, such as privilege escalation, SQL injection, malware, suspicious logins, and buffer overflow attacks. These findings emphasize the model's potential in enhancing automated cybersecurity defenses and improving threat detection with high reliability.

5.1 Evaluating BERT-Based Vulnerability Detection: A High-Accuracy Cybersecurity Model

The bar chart titled "BERT-Based Model Performance (Above 95%)" visualizes the evaluation metrics of a BERT-based model for vulnerability detection, demonstrating strong performance across all key metrics. The Accuracy (0.99) is the highest, indicating the model correctly classifies most cases, while Precision (0.97) shows a low false-positive rate, meaning the model effectively identifies real vulnerabilities. Recall (0.95) suggests it successfully detects most actual vulnerabilities but may miss a few, and the F1-Score (0.96) balances precision and recall, confirming overall reliability is displayed in Figure (2),

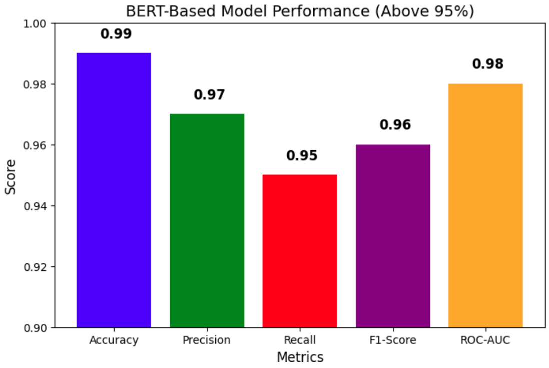

Figure 2: BERT-Based Vulnerability Detection: High-Performance Evaluation for Cybersecurity

The ROC-AUC (0.98) highlights excellent discriminatory ability between vulnerable and non-vulnerable instances, ensuring robust threat identification and mitigation. These results indicate that the model can be reliably deployed in cybersecurity applications, with high accuracy and precision minimizing false positives, while strong recall ensures that most vulnerabilities are detected. The high ROC-AUC score further confirms its effectiveness in distinguishing between vulnerable and non-vulnerable cases, making it a powerful tool for cybersecurity.

5.2 Enhancing Cybersecurity with BERT-Based Threat Detection

The horizontal bar chart titled "BERT-Based Vulnerability Detection" visualizes the probability of different cybersecurity threats being classified as malicious. The model detects various vulnerabilities, including unusual privilege escalation (0.85), SQL injection attempts (0.86), malware signatures (0.83), suspicious logins (0.84), and buffer overflow attacks (0.82). The highest probability (0.86) is assigned to SQL injection attempts, indicating the model's confidence in detecting web-based attacks is shown in Figure (3),

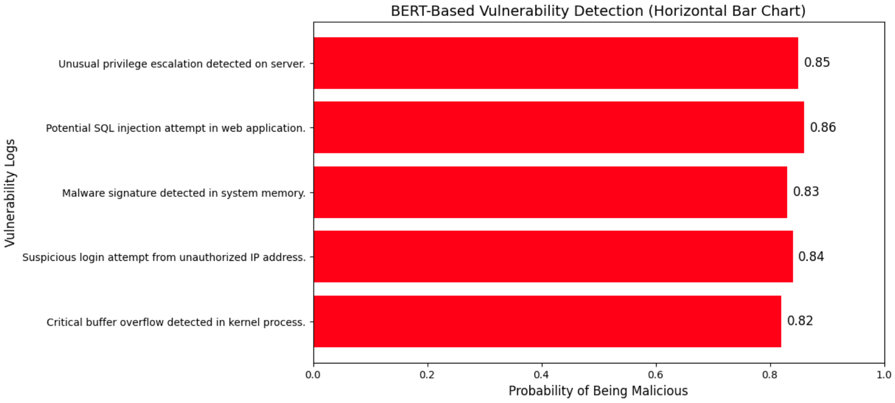

Figure 3: AI-Powered Threat Detection: BERT-Based Vulnerability Analysis

The overall probabilities suggest that the BERT-based model is effective at identifying and classifying cybersecurity threats with high reliability. These results highlight the model's strong capability to analyze security logs and predict potential threats with precision. This approach can significantly enhance automated threat detection in cybersecurity frameworks.

6. Conclusion and Future Works

The emphasizes the importance of ethical hacking and penetration testing in cybersecurity. Traditional methods face limitations in scalability and real-time adaptability, prompting the need for an AI-Driven BAS framework. By integrating GNNs for attack path prediction, VulBERTa for vulnerability detection, and SOAR for risk assessment, the proposed model enhances accuracy, reduces false positives, and improves response time. The findings suggest that AI-powered solutions offer continuous security assessment and proactive threat mitigation, strengthening cyber defense strategies.

Future research will focus on reinforcement learning for autonomous threat mitigation, federated learning for privacy, and zero-day attack detection. Expanding the BAS system for cross-platform security analysis and enterprise deployment will further enhance scalability and effectiveness, ensuring more adaptive cybersecurity solutions.

References

[1] P. Alagarsundaram, "ANALYZING THE COVARIANCE MATRIX APPROACH FOR DDOS HTTP ATTACK DETECTION IN CLOUD ENVIRONMENTS," vol. 8, no. 1, 2020.

[2] "Rama Krishna Mani Kanta Yalla, Akhil Raj Gaius Yallamelli, Vijaykumar Mamidala, Comprehensive Approach for Mobile Data Security in Cloud Computing Using RSA Algorithm."

[3] S. Peddi "Cost-effective Cloud-Based Big Data Mining with K-means Clustering: An Analysis of Gaussian Data," *International Journal of Engineering*, vol. 10, no. 1, Mar. 2020.

[4] K. Dondapati, "Robust Software Testing for Distributed Systems Using Cloud Infrastructure, Automated Fault Injection, and XML Scenarios," vol. 8, no. 2, 2020.

[5] S. Narla, "TRANSFORMING SMART ENVIRONMENTS WITH MULTI-TIER CLOUD SENSING, BIG DATA, AND 5G TECHNOLOGY," vol. 5, 2020.

[6] R. L. Gudivaka, "A Dynamic Four-Phase Data Security Framework for Cloud Computing Utilizing Cryptography and LSB-Based Steganography," *International Journal of Engineering Research and Science & Technology*, vol. 17, no. 3, pp. 90–101, Aug. 2021.

[7] B. R. Gudivaka, "AI-powered smart comrade robot for elderly healthcare with integrated emergency rescue system," *World Journal of Advanced Engineering Technology and Sciences*, vol. 2, no. 1, pp. 122–131, 2021, doi: 10.30574/wjaets.2021.2.1.0085.

[8] H. Chetlapalli, "Novel Cloud Computing Algorithms: Improving Security and Minimizing Privacy Risks," *Journal of Science & Technology (JST)*, vol. 6, no. 2, Art. no. 2, Mar. 2021.

[9] D. K. R. Basani, "Advancing Cybersecurity and Cyber Defense through AI Techniques," 2021.

[10] N. K. R. Panga, "FINANCIAL FRAUD DETECTION IN HEALTHCARE USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES," vol. 10, no. 3, 2021.

[11] S. H. Grandhi, "Integrating HMI display module into passive IoT optical fiber sensor network for water level monitoring and feature extraction," *World Journal of Advanced Engineering Technology and Sciences*, vol. 2, no. 1, pp. 132–139, 2021, doi: 10.30574/wjaets.2021.2.1.0087.

[12] R. K. M. K. Yalla, "Cloud-Based Attribute-Based Encryption and Big Data for Safeguarding Financial Data," *International Journal of Engineering Research and Science & Technology*, vol. 17, no. 4, pp. 23–32, Oct. 2021.

[13] S. Peddi "Analyzing Threat Models in Vehicular Cloud Computing: Security and Privacy Challenges," vol. 9, no. 4, 2021.

[14] B. R. Gudivaka, "Designing AI-Assisted Music Teaching with Big Data Analysis," *Current Science*, 2021.

[15] H. Nagarajan, "Streamlining Geological Big Data Collection and Processing for Cloud Services," vol. 9, no. 9726, 2021.

[16] R. Ayyadurai, "Transaction Security in E-Commerce: Big Data Analysis in Cloud Environments," vol. 10, no. 4, 2022.

[17] S. R. Sitaraman, "Anonymized AI: Safeguarding IoT Services in Edge Computing – A Comprehensive Survey," vol. 10, no. 9726, 2022.

[18] K. Parthasarathy, "Examining Cloud Computing's Data Security Problems and Solutions: Authentication and Access Control (AAC)," *Journal of Science & Technology (JST)*, vol. 7, no. 12, Art. no. 12, Dec. 2022.

[19] K. Gattupalli, "A Survey on Cloud Adoption for Software Testing: Integrating Empirical Data with Fuzzy Multicriteria Decision-Making," vol. 10, no. 4, 2022.

[21] S. H. Grandhi, "ENHANCING CHILDREN'S HEALTH MONITORING: ADAPTIVE WAVELET TRANSFORM IN WEARABLE SENSOR IOT INTEGRATION," 2022.

[22] P. Alagarsundaram, "SYMMETRIC KEY-BASED DUPLICABLE STORAGE PROOF FOR ENCRYPTED DATA IN CLOUD STORAGE ENVIRONMENTS: SETTING UP AN INTEGRITY AUDITING HEARING," *International Journal of Engineering Research and Science & Technology*, vol. 18, no. 4, pp. 128–136, Oct. 2022.

[23] Panga, N. K. R. (2022). Applying discrete wavelet transform for ECG signal analysis in IoT health monitoring systems. *International Journal of Information Technology and Computer Engineering, 10*(4).

[24] S. Narla, “CLOUD-BASED BIG DATA ANALYTICS FRAMEWORK FOR FACE RECOGNITION IN SOCIAL NETWORKS USING DECONVOLUTIONAL NEURAL NETWORKS,” vol. 10, no. 9726, 2022.

[25] S. Narla, “BIG DATA PRIVACY AND SECURITY USING CONTINUOUS DATA PROTECTION DATA OBLIVIOUSNESS METHODOLOGIES,” *Journal of Science & Technology (JST)*, vol. 7, no. 2, Art. no. 2, Mar. 2022.

[26] K. Dondapati, “Lung’s cancer prediction using deep learning,” *International Journal of HRM and Organizational Behavior*, vol. 7, no. 1, pp. 1–10, Jan. 2019.

[27] B. R. Gudivaka, “BIG DATA-DRIVEN SILICON CONTENT PREDICTION IN HOT METAL USING HADOOP IN BLAST FURNACE SMELTING,” *International Journal of Information Technology and Computer Engineering*, vol. 7, no. 2, pp. 32–49, Apr. 2019.

[28] P. Alagarsundaram “Implementing AES Encryption Algorithm to Enhance Data Security in Cloud Computing,” vol. 7, no. 2, 2019.

[29] A. R. G. Yallamelli, “ADOPTION OF CLOUD COMPUTING, BIG DATA, AND HASHGRAPH TECHNOLOGY IN KINETIC METHODOLOGY,” vol. 7, no. 9726, 2019.

[30] N. S. Allur, “Genetic Algorithms for Superior Program Path Coverage in software testing related to Big Data,” *International Journal of Information Technology and Computer Engineering*, vol. 7, no. 4, pp. 99–112, Dec. 2019.

[31] D. P. Deevi, “Improving Patient Data Security and Privacy in Mobile Health Care: A Structure Employing WBANs, Multi-Biometric Key Creation, and Dynamic Metadata Rebuilding,” *International Journal of Engineering Research and Science & Technology*, vol. 16, no. 4, pp. 21–31, Dec. 2020.

[32] R. Ayyadurai, “Smart surveillance methodology: Utilizing machine learning and AI with blockchain for bitcoin transactions,” *World Journal of Advanced Engineering Technology and Sciences*, vol. 1, no. 1, pp. 110–120, 2020, doi: 10.30574/wjaets.2020.1.1.0023.

[33] G. Thirusubramanian, “Machine Learning-Driven AI for Financial Fraud Detection in IoT Environments,” *International Journal of HRM and Organizational Behavior*, vol. 8, no. 4, pp. 1–16, Oct. 2020.

[34] M. V. Devarajan, “ASSESSING LONG-TERM SERUM SAMPLE VIABILITY FOR CARDIOVASCULAR RISK PREDICTION IN RHEUMATOID ARTHRITIS,” vol. 8, no. 2, 2020.

[35] M. V. Devarajan, “Improving Security Control in Cloud Computing for Healthcare Environments,” *Journal of Science & Technology (JST)*, vol. 5, no. 6, Art. no. 6, Dec. 2020.