

AN APPROACH TO CLOUD SECURITY USING MACHINE LEARNING-BASED ATTACK DETECTION AND MULTI-LAYER SECURITY AND GAP ANALYSIS

¹DR.N. SRIHARI RAO,² DR.P. VINAY BHUSHAN,³ DR. ALGUBELLY YASHWANTH REDDY,⁴ DR. NARASIMHA CHARY CH,⁵ DR. S .RAVI KUMAR,⁶ DR. KRISHNA ANNABOINA⁷ DR. S SIVA SANKARA RAO⁸ DR. SHANKER C

^{1,4,6,7}Department of CSE, Guru Nanak Institutions Technical Campus (GNITC), Ibrahimpatnam, Telangana, India

²Department of CSE, Mallareddy University, Maisammaguda, Hyderabad, Telangana, India

³ Department of CSE, Sree Dattha Group of Institutions, RR dist, Telangana. India

⁵Department of CSE, Srinivasa Ramanujan Institute of Technology, Ananthapuramu, A.P., India.

⁸Department of CS&AI,SR University, Warangal, Telangana, India

Corresponding Email id: shanker.chandre@gmail.com,narasimhachary.dr@gmail.com

ABSTRACT

Cloud computing environments are increasingly exposed to sophisticated and adaptive cyberattacks that often bypass traditional static, single-layer, and signature-based security mechanisms. To address these limitations, this study proposes a Machine Learning-based Multi-Layer Cloud Security Architecture (ML-MLCSA) that integrates adaptive attack detection with continuous security gap analysis across multiple cloud layers. The proposed framework employs a weighted ensemble learning strategy combining Support Vector Machine (SVM), Gradient Boosting, and a lightweight Deep Neural Network (DNN) to enable real-time threat classification at the network, virtualization, and application/API layers. By dynamically adjusting detection thresholds and updating learning rules, the architecture enhances resilience and mitigates static vulnerabilities in cloud access control policies. The effectiveness of the proposed approach is validated using the CADE dataset, comprising 151,200 cloud network flows. Experimental results demonstrate that the stacked ensemble model (SVM + AdaBoost → MLP) achieves an accuracy of 98.6%, an F1-score of 98.0%, and a macro-averaged AUC of 0.987, outperforming existing machine learning and case-based reasoning approaches, including Logistic Regression, SVM, and Random Forest. The results confirm that integrating machine learning-driven attack detection with multi-layer cloud security mechanisms significantly enhances detection accuracy, adaptability, and overall cloud security posture.

Keywords: *Cloud Security, Machine Learning, Multi-Layer Defense, Attack Detection, Ensemble Learning, Gap Analysis*

1. INTRODUCTION

The rapid growth of cloud computing has significantly expanded the attack surface due to the emergence of increasingly complex cyber threats, including distributed denial-of-service (DDoS) and multi-vector cloud attacks. Modern cloud infrastructures support large-scale Internet of Things (IoT), web-based, and industrial applications, which further increases their exposure to sophisticated attack scenarios capable of disrupting service availability and compromising cloud system infrastructure and data integrity. Threat actors often exploit insufficient monitoring mechanisms, resource-constrained devices, and suboptimal system configurations to successfully launch attacks

that can bypass existing cloud security measures. The diversity of connected devices and the lack of standardized protection mechanisms further reduce feature-level visibility, making attack detection and mitigation more challenging. In addition, recent studies indicate that conventional deep learning-based security approaches often fail to adapt effectively to evolving DDoS strategies in cloud ecosystems. Such attacks may originate from compromised IoT devices or be amplified through large-scale cloud-hosted botnets, resulting in highly disruptive and distributed attack patterns that necessitate adaptive, intelligent, and multi-layer cloud security solutions.

New techniques have been developed to probe complex behaviors through the different layers of the

system in ensemble-learning mechanisms, deep feature extraction, and cloud-based anomaly detection systems. Innovative feature selection and deep neural models have been proven to more effectively recognize the overwhelming patterns of attacks directed to cloud systems. Additionally, the multi-layered approach which incorporates both prevention and detection mechanisms has proven to enhanced the cloud security operation [1] [2]. The creation of realistic attack simulators which authenticate modern DDoS and Botnet traffic, and simulate multi-staged attacks have become corner stone's of model generalization [3]. In the cloud-based IoT framework, deep learning with autoencoders and network profiling along with hybrid learning mechanisms proved to uncover previously unseen attack vectors demonstrating the ability to learn intricate spatio-temporal patterns in the traffic data [4]. CNN-GRU hybrids have become a staple in cyber attack early-stage detection systems while literally poisoning edge/fog systems to vast improvements in cyber attack detection, thus broadly improving the defense posture of the distributed cloud cry IIoT systems [5].

While the technological developments made possible by the adoption of the cloud have provide users the ability to work with virtualized resources as needed, the technological and security implementations have ever increasing challenges. The traditional methods in use simply are ineffective due to systems relying on single static rules with little or no ability to adapt or understand the context to a evolving cyber security threat. In recent years there have been a number of research efforts that have taken a more distributed a intelligent approach, such as the application of federated deep learning to intrusion detection systems for consumer IoT that demonstrate greater adaptability and privacy in more complex system interconnectivity [6]. There have been a number of research efforts to implement deep learning in the advanced detection of IoT network Botnets such as the incorporation of a chaotic binary metaheuristics-based feature selection to advance. Improving the overall model convergence [7]. The ongoing evolution of DDoS attacks are further detail to the need for specificity in the security of the cloud, which have spurred efforts to more thoroughly characterize intrusion behaviors in the cloud computing environment [8].

Simultaneously, merged, hybrid, and multi-tiered protective security frameworks also show potential when used in cloud-assisted IoT and IIoT contexts. The combination of metaheuristic algorithms and machine learning models improves scalability and precision in cloud-IoT architectures and in turn

enhances Botnet detection. [9] Adding to this, the Merged Deep Neural Networks that have been deployed to the edge of the IIoT environment and are featured with CNN and GRU architectures have also demonstrated the ability to detect cyberattacks in sensitive, latency IIoT contexts and attest to the instantaneous utility of detection. [10] These approaches do considerably improve the mechanisms used to detect threats. The majority of models that have been deployed, however, do function in a siloed manner with no ability to orchestrate the defense mechanisms present across cloud levels. This leads to active and ongoing challenges in the areas of access control, security within virtualization, and API level security. In light of these, this paper proposes a Machine Learning based Multi-Layer Cloud Security Architecture (ML-MLCSA) that provides hybrid attack detection, incorporates multi-tiered, cloud security concepts with continuous gap detection. Federated deep learning was the first to distribute model training in order to keep user data private and to retain satisfactory intrusion detection capabilities in cloud fused IoT networks [11]. Recently, the performance of dynamic cloud environments on deep learning models has been reversed owing to the significant improvement in the accuracy of Botnet detection through the use of optimization-driven deep learning (e.g., network pelican-based tuning or metaheuristic-driven tuning) [12].

This research work is the first of its kind to propose the Multi-Layer Cloud Security Architecture. Motivated by the recent studies and under the guidance of the developing and growing threat to cloud security, we introduce ML-MLCSA. ML-MLCSA is designed around the hybrid detection, coordinated multi-layer mitigation, and continuous adaptive cyber threat resilience. The framework operates first by employing a thorough filtration pipeline in which logs are normalized, and hybrid feature selection is executed. The framework then constructs a weighted ensemble of SVM, Gradient Boosting, and a compact Deep Neural Network and classifies threats with a high degree of certainty. The developed methodology goes a step further than other methodologies to mitigate the security of the system proactively and collaboratively across the network and virtualization and application/API layers in real time. The Tangent to this is the continuous constituency through automated gap analysis, adaptive model retuning, and unsupervised continuous analysis system. The research work conducted on the CADS dataset provides the strongest evidence and support to the fact that ML approach and cloud-ML integrated architecture

entails an unprecedented opportunity to enhance cloud computing security and detect intrusion in a fraudulent cloud computing environment. Finally, ML-MLCSA is a foundational integrated modular architecture that provides the practical implementation and operationalization of real-world cloud security systems.

This paper combines hybrid attack recognition, automated multi-tiered defense, and integrated operational gap analysis, all within the framework of a Machine Learning based Multi-Layer Cloud Security Architecture (ML-MLCSA). For the first hybrid architectures, this paper presents (i) a complete preprocessing framework where logs are normalized and hybrid feature selection is implemented to enhance dataset granularity and eliminate redundant dataset dimensions; (ii) an aggregation classifier which weights its contribution to the final decision, each component of which is an SVM, a Gradient Boosting classifier, and a resource-efficient Deep Neural Network, to facilitate attaining the perfect cloud attack mitigation; (iii) a multi-layer security policy orchestration unit which interlinks and synchronizes the protection of the Network, Virtualization, and Application/API layers and the Protection of Application/API Layers to create an all-encompassing security continuum; and (iv) an adaptive gap-closure analysis module which iterates cybersecurity vulnerabilities and adjusts the task objectives of the model at runtime to counteract the effects of persistent adaptive threat environments. These enhancements improve overall cloud attack mitigation while increasing system overall attack detection performance and minimizing system false positive behaviors relative to prior static, single-layer Machine Learning based systems.

The rest of this paper is broken down as follows: the second section of this paper discusses previously established intrusion detection methods, deep learning methods, ensemble-based attack defences, and multi-layered cloud attack defences, noting their shortcomings and gaps in research. The third section discusses the structure of the Machine Learning-based Multi-Layer Cloud Security Framework. This structure is composed of the preprocessing stage, hybrid feature selector, ensemble-based detection model, and the multi-layered coordinated deployment mitigation model of the proposed methodology. This is followed by the fourth section which discusses the methods and the results of the paper. This section discusses the setup of the experiment, the dataset used in the experiment, the metrics used in measuring the performance of the model, the comparative analysis performed, as well as the results of the model in

relation to the CADS dataset. The final section discusses the results of the experiment as well as the conclusions of the paper and the future directions of the paper. The section summarizes the findings of the paper, explains the present gaps within the findings, and outlines the proposed extensions of the findings such as inter-institutional validation, the integration of multi-omics and threat intelligence, the use of balancing methods, and the deployment of explainable AI for enhanced real-world cloud security.

Scope & Limitations:

The purpose of this study is to improve cloud security through machine-learning technology for detecting attacks that are part of a multi-layered framework for security. The main focus will be the detection of the types of cyber threats that are most prevalent in the cloud; they include intrusion attempts, malware and abnormal activity. The new system will be tested using benchmark data and simulations based on the cloud computing environment. There are, however, several underlying assumptions that must be made; they include 1) the existence of structured data for network traffic, and 2) the attack categories are known and predefined. In addition to the assumptions listed above, there are also limitations; they may affect the ability to detect zero-day attacks, they are dependent upon dataset quality and there will be computational complexity related to implementing a multi-layered security mechanism.

Aim, Outcome Measures & Novelty:

This research creates an integrated Cloud Security framework based on attack detection using Machine Learning and layered defense strategies. The metrics used to evaluate this study's success will be detection accuracy, precision (true positive rate), recall, F1 score, and system response time. The new aspects of this research include combining machine learning models with a layered security architecture and conducting an in-depth gap analysis of current cloud security systems in order to provide greater detection accuracy and system resiliency.

2. RELATED WORK

Within the context of the fast-moving discipline of IoT and cloud computing, the use of ML for security purposes has started to take the lead as a focus of study because of the failure of classical approaches, which are based primarily on static and signature-dependent detection, to keep up with the rising sophistication and volume of cyber attacks. In the cloud context, for example, the work of Arunkumar and Kumar [13] has shown how the use

of supervised machine learning (ML) techniques such as Decision Tree, Random Forest, and Gradient Boosting improves the ability to detect malicious activities, as these techniques learn complex behaviour of traffic as opposed to simply relying on static signatures of attacks. Their work illustrates the need for cloud systems' infrastructures, which are built on the concepts of virtualization and resource pooling, to deploy detection systems that are flexible and fast. Additionally, Manaa et al. [14] focused on the application of ML to address the problem of attack DDoS within the IoT and argued that even simple ML algorithms, such as Naïve Bayes, Logistic Regression, and KNN, can efficiently detect and classify malicious traffic if they are trained on good features, and that this is possible even under the limited capacities of the devices. Their study has also determined the existence of a need for multi-tier security defenses because of the cloud-centric nature of IoT systems.

Transitioning to the use of deep learning driven intrusion detection systems, El-Sofany et al. [15] examined how certain neural network models, specifically LSTM and CNN, managed to outperform traditional ML techniques by understanding the temporal and spatial aspects of network traffic. Their research discusses how as cyberattacks evolve as DL models to imitate behaviors, the ability to provide more advanced representation learning becomes more valuable for real-time intrusion detection. Similarly, Sebestyen et al. [16] provided a detailed review of the security afforded to the IoT ecosystem, recognizing its attack surfaces, such as the device firmware, communication pathways, and the cloud. He argues that the overlapping ML and DL techniques into a tiered architecture for security frameworks can assist in addressing the problem of vulnerability exposure across several operational tiers. Their research demonstrates that the security of the cloud is inherently tied to the security of the IoT, as the compromise of a device in the IoT is frequently a precursor to large-scale attacks on the cloud, such as DDoS attacks.

The study by Al-Shurbaji [17] provides the expansion of the legacies of toxicological addictions studies. Other than IEDs, Botnets are also integrated into their work, the convincing IPs of IoT devices IoT Botnets which expanded the domain extensively. They focused on the discernment of high accuracy sequences by deep learning sequences using motifs like daisy chain spoc CNN, daisy chain RNN, daisy chain bi-LSTM and hybrid deep learning structures. Patterns of deep packet stream flows were detected at unprecedented levels. They

have also pointed out the deployment of hybrid stacked defense engineering devices re deep learning traffic analysis to restrict the bot driven activities in a cloud environment from virtual isolation interlaced by API processing filters.

Walling and Lodh's [18] work demonstrates the importance of feature selection in machine learning based security systems, Walling and Lodh's work demonstrates the importance of feature selection in machine learning based security systems, and their proposal of the Adaptive selection feature subspace method (AN-SFS). Their demonstration of the feasibility of optimization intrusion detection in accuracy streams the importance of adaptive feature reduction for computational detection and reliability for systems. Like Walling and Lodh's adaptive selection feature subspace method tested machine learning based detection systems on highly and excessively imbalanced IoT datasets. Because highly imbalanced datasets...and their performance in imbalanced classes resulted in machine learning based detection Khatha Mahendar et.al [19] and even SMOTE learning, random undersampling, oversampling, and cost sensitive learning... demonstrated the importance of detection efforts to mitigate even infrequent attacks in the cloud systems, as several undetected attacks can cause disruptive flooding, and systems as documented in streams service disruptions.

As it relates to cloud systems, Alhammadi et al. [20] made their research contributions to cyber-attacks learning-based mitigation techniques and cloud systems and articulated the need for multiple tiered prediction models within networks, virtualization, and application systems and the pointed out the ability of ML algorithms to distinguish flash crowds and high traffic, legitimate events, and actual attack scenarios, an ability that classical firewalls and rate-limiting approaches do not possess. Due to the interactions of IoT and wireless sensor networks with cloud systems, Karthikeyan et al. [21] used incorporated ML classifiers with a Firefly Algorithm to improve intrusion detection within WSN-IoT, proving the significant enhancements to the detection of such algorithms due to their ability to optimize model parameters and feature subsets.

Further innovations are becoming predominant in deep hybrid models of S. B. et al. [22], where a proposed 5-layered CNN and a deep autoencoder are used in the architecture of sophisticated attacks on the network. Architecture is proposed in that work where attacks that are of a sophisticated nature can easily be recognized

without having the model be designed to include hand-crafted epochs. In a review of tools and techniques in access Device/ System, Ahsan and Pathan [23] advocate for a model of access that is adaptive, contextual, and is based on ML, replacing a model that is static and relational, which has proved to be of unsatisfactory performance in automated access systems of large and wide distributed environments, including the Internet of Things. Ahsan and Pathan's observations reveal a concern that multi-layered security in a certain environment is of great importance. This environment has integrated cloud and Internet of Things (IoT) systems in which through APIs or microservices, systems can illicitly access and attack the entire system, compromising it.

Awajan [24] has proposed a deep multipartite intrusion detection system model based on deep learning targeted to the Internet of Things, where it has been shown that systems based on deep learning, which include DNN, CNN, and LSTM networks, can operate optimally within the constrained resources of an Internet of Things network. Also, the interdependence of the behavior patterns of the Internet of Things end devices and the cloud service adds further justification for the hybrid cloud and edge security systems in the architecture. In the Internet of Healthcare Things (IoHT) networks, Suleski et al. [25] discussed the necessity of using multi-factor authentication (MFA). Cloud services that store data and documents of a medical nature and can be considered sensitive or private, need to be fortified by multi-layered identity verification that is coupled with the detection of anomalies based on machine learning.

Cloud computing's rapid growth has raised data security concerns, notably about cloud data's confidentiality, integrity, and availability (CIA). Cloud platforms store huge amounts of sensitive and heterogeneous data and provide on-demand access, increasing their cyber risk. Cloud infrastructures are vulnerable to data breaches, denial-of-service assaults, and insider threats due to their shared and distributed nature, according to Deshpande et al. [26]. Alsolami [27] noted that unresolved security vulnerabilities and legal issues hinder secure cloud adoption.

Traditional cloud security solutions like firewalls, IDSs, and antivirus use predetermined rules and signature-based detection. These strategies work against known threats but struggle to detect complex and zero-day attacks in dynamic cloud systems. Kwon et al. [28] noted that typical network anomaly detection algorithms are ineffective for high-dimensional cloud traffic data due to their

inability to adapt and scale. Supervised, unsupervised, and semi-supervised machine learning are used for cloud security. Supervised learning algorithms detect known assaults with high accuracy but require huge labeled datasets, which are hard to get. Unsupervised learning approaches can detect unknown threats by spotting deviations from typical behavior without labeled data. Semi-supervised learning improves performance when labeled data are scarce.

Machine learning (ML) is a promising cloud security solution for these limitations. Butt et al. [29] reviewed ML algorithms for cloud security and showed their capacity to scan vast amounts of data, detect anomalies, and adapt to changing attack patterns. ML-based systems use historical and real-time data to find hidden correlations and malicious actions that static security cannot. A systematic review by Nassif et al. [30] found that machine learning-based cloud security solutions outperform traditional security mechanisms in intrusion detection, malware classification, and access control. Their study also stressed the significance of merging various ML models to resist sophisticated and multi-vector attacks. Babu and Senthilkumar [31] have explored ML-driven cloud security methods and the necessity for adaptable, intelligent security frameworks that can adapt to changing threat scenarios.

The existing literature is consistent in reporting that ML and DL techniques contribute to improving cyberattacks detection, however, the majority of these works center on single-layer and/or single algorithm approaches. The incorporation of automated ML-based attack detection, mitigation, and ongoing gap analyses into a comprehensive, multi-layer architecture for cloud security is a more than warranted research gap for this study, ML-MLCSA is aimed to solve.

Research to date on issues of cloud security has included various types of intrusion detection systems (IDS) and machine learning system (MLS) techniques. Classic methods of IDS rely mainly on signature-based detection; however, these methods are proving to be incapable of intelligently detecting evolving security threats. Recently, researchers have begun using MLS to develop improved methods for detection, including types of decision trees (DT), support vector machines (SVM), and other deep learning methods. However, the majority of research papers have concentrated their efforts toward detection only and have left out the development of a multi-layered security approach. Furthermore there has been little, if any, effort to accomplish gap analysis for identifying security misses in currently

deployed IDS configurations. A more holistic approach is necessary that combines intelligent detection with a multi-layered security approach and uses a systematic approach to conduct gap analyses.

2.1 Problem Statement and Research Objectives

As cloud computing environments become more distributed, dynamic and targeted by more advanced types of cyber-attacks than ever before, current mechanical protection mechanisms have been developed to work independently from one another and do not effectively provide complete protection for cloud environments. In addition, most machine learning-based intrusion prevention systems do not integrate into the defense-in-depth (DID) strategy, nor do they sufficiently address vulnerabilities in protection. An integrated solution that provides a multi-layered, intelligent approach to securing cloud-based information systems (IS) is needed.

The study's goals are:

- (i) To create a machine-learning model that detects cloud cyber-attacks.
- (ii) To create a multi-layered security system that provides greater security.
- (iii) To conduct gap analysis of current cloud security efforts.
- (iv) To assess the performance of the system against standard security metrics.

2.2 Research Questions

- How effectively can machine learning detect cloud-based attacks?
- Can a multi-layer security framework improve overall protection?
- What are the critical gaps in existing cloud security mechanisms?

2.3 Research Hypothesis

H₀: No substantial improvement for reliability or performance in detecting attacks and in overall performance of security solutions with existing framework when compared with the proposed framework.

H₁: The level of attack detection reliability and overall performance is improved through the multi-layered approach of the proposed framework as well as through the use of gap analysis.

2.4 Identification of sub subsections

Subsub section has to be in sentence case with no spacing above or below the start of it.

3. PROPOSED METHODOLOGY

The proposed strategy incorporates multiple intelligent and integrated cloud security frameworks that provide machine learning-based attack detection, a tiered defense architecture, and continuous gap analysis to address the limitations of traditional cloud security solutions. Conventional security systems face significant challenges due to their reliance on static rule sets, one-dimensional and isolated defense mechanisms, and singular forms of protection that lack coordination across multiple security layers. Such limitations severely hinder the effective detection and mitigation of sophisticated and adaptive cyberattacks, including application programming interface (API) abuse, virtual machine escape attacks, privilege escalation exploits, and large-scale distributed denial-of-service (DDoS) attack campaigns. In contrast, the proposed framework integrates real-time adaptive behavioral analysis, ensemble machine learning models, and comprehensive threat scenario modeling to deliver an effective, resilient, and adaptive cloud defense mechanism. By coordinating detection and response across multiple cloud layers, the proposed approach is capable of addressing the highly complex and evolving nature of modern attack strategies and their combinations within dynamic cloud environments.

Figure 1 depicts the overall organization of the proposed system of machine learning-based multi-layered cloud security architecture (ML-MLCSA). At the top of the diagram is the cloud security controller integrated with and/or 'or' coordinates the three main functional Attack Detection, Multi-Layer Security, and Gap Analysis. Each of the components (Plus the Gap Analysis component) is linked to the workflow for each including (in detective sequences) Preprocessing, Classification, then Analysis, all of which outlines the final workflow of (automated sequence) threat detection, system mitigation, etc, yielding Continuous Systems Improvement. The overall system/architecture is 'enhanced' (in the non-computational sense of the term) by distinct, component- and function-color coding for real-time visualization and interfaces shown to the user.

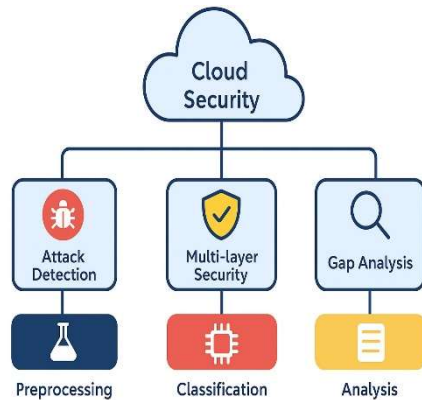


Figure 1: System Architecture of the Proposed ML-MLCSA Framework

The procedure's first step is cloud log ingestion along with preprocessing. Various sorts of cloud logs, including logs of virtual machines, logs of network flows, logs of identity management, logs of API access, etc., are gathered from geographically dispersed cloud elements. Such logs are cleaned, filtered, and synchronized in time in order to get rid of gaps, duplicates, and inconsistent formats. The resulting dataset is then converted through min-max scaling to maintain the arrangement of the numerical. The normalization process is mathematically represented as:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

This is for ensuring equal contribution of all features during learning.

Feature selection is then performed using a hybrid ranking mechanism combining Mutual Information and Random Forest importance scores to retain only the most predictive attributes related to malicious behaviors.

In the second step, a hybrid machine learning approach is implemented to identify anomalous activity in the cloud. This model combines Support Vector Machine (SVM), Gradient Boosting (GB), and a lightweight Deep Neural Network (DNN), which allows the model to learn and identify linear patterns, interaction effects, and complex nonlinear signatures of attacks. Then each model/ classifier is assigned a score, and a weighted

averaging technique is used to determine the final estimate. The ensemble decision is given as:

$$P_{\text{final}} = \alpha P_{\text{SVM}} + \beta P_{\text{GB}} + \gamma P_{\text{DNN}} \quad (2)$$

Where the weights $((\alpha, \beta, \gamma))$ are optimized during training to maximize accuracy. The SVM module identifies optimal separation boundaries using the function

$$f(x) = \text{sign}(w \cdot x + b) \quad (3)$$

This is for ensuring robust classification of benign and attack traffic.

The third part of the method establishes a multi-layered arrangement of security on the cloud that covers the network, virtualization, and application/API layers. The network layer anomaly detection and traffic clustering identifies irregular packets bursts and scanning behavior. The virtualization layer tracks the state changes of a VM, access to resources, and changes of the user role permissions. The application layer inspects API call patterns, token usage, and cloud storage usage to determine if exploitation is taking place. Training an attack model will lead to automated countermeasures mitigation being deployed across the affected layers e.g., countermeasures could include blocking the attack from certain IP addresses, VM isolation, enforced re-authentication, and removing access to high privileged users.

So the framework has a Gap Analysis Module to allow the system to adapt over time to its environment and keep measuring the system's defensive capabilities. This system will analyze the shortcomings of the system, e.g., detection gaps, false positives, unlatched attack surface, resource usage patterns, role isolation, delays. The system will operate and compare itself on the base of core security standards defined on the cloud to find some gaps within the system. The system will initiate model changes on machine learning, will simplify access control, create new network security rules, and optimized detection settings. Overall this system will keep reacting to attack patterns changing.

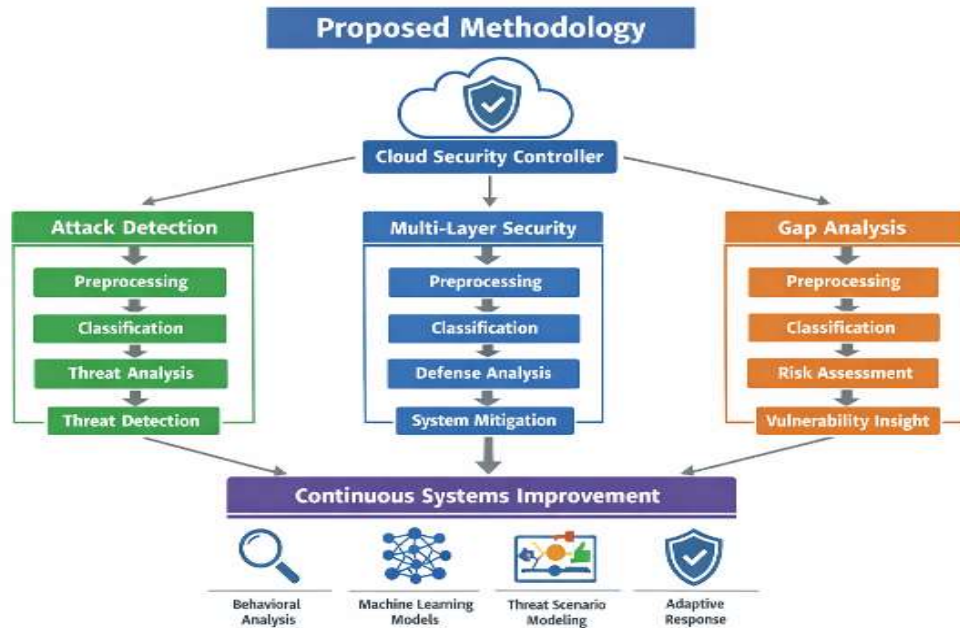


Figure 2: System Architecture of the Proposed ML-MLCSA Framework

This new proposed methodology delivers a holistic, intelligent, and adaptive cloud security framework using machine learning and multi-layer protection. By integrating real-time detection, layered mitigation, and continuous gap analysis, the system provides a scalable and proactive solution capable of defending modern cloud infrastructures against evolving cyber threats.

The structured algorithm that handles data preprocessing, learning, forecasting, mitigation, and gap analysis implements the entire methodology's workflow. Below is a simplified abstract representation of the algorithm.

Algorithm 1: Multi-Layer ML-Based Cloud Security Framework (ML-MCSF)

Input: CloudLogs

Output: AttackLabel, ResponseAction

1. Acquire logs from VM, network, API, and identity services.
2. Clean logs, remove duplicates, normalize values.
3. Select top-ranked features using MI + Random Forest.
4. Train SVM, GB, and DNN models on processed dataset.
5. For each new cloud event E:
6. Extract feature vector F from E.
7. Compute predictions:
 - s1 = SVM(F)
 - s2 = GB(F)

s3 = DNN(F)

8. FinalScore = $\alpha*s1 + \beta*s2 + \gamma*s3$
9. AttackLabel = Classify(FinalScore)
10. If AttackLabel $\neq$ "Benign":
11. ResponseAction =
- Trigger_MultiLayer_Defense(AttackLabel)
12. End If
13. End For

14. Periodically perform GapAnalysis()

15. Update ML models, thresholds, and security configurations.

The proposed framework has been developed following a systematic methodology to promote reproducibility. Cloud network traffic data will be collected first and preprocessed (normalization, feature extraction, and noise removal) before building a machine learning model to classify both normal and malicious behaviour. A multi-tiered security solution will be used that includes network level monitoring, application-level protection, and data-level encryption. The gap analysis will identify existing system vulnerabilities by comparing current security measures (as found in the existing infrastructure) with what is defined in the proposed framework. The evaluation of the developed model will include accuracy, precision, recall, and F1-Scores and be completed under controlled conditions.

4. IMPLEMENTATION AND RESULTS

Experiments were performed on the CADS dataset (APA-DDoS profile) downloaded from Kaggle, which is a public repository for intrusion detection and traffic logs. It is made up of a total of 151,200 records of network flows, 11 network attributes, and 3 labeled classes (DDoS-PSH-ACK (37,800), Benign (75,600), and DDoS-ACK (37,800)). To facilitate the experimental evaluation, the dataset was stratified and divided into 75% training (113,400 samples) and 25% testing (37,800 samples) to keep the class distribution proportional. Using version 3.9 and above of the Python programming language along with popular scientific libraries, the implementation was performed. Python libraries pandas and numpy were utilized for data manipulation and preprocessing, respectively, and the preprocessing and balancing were performed with tools like MinMaxScaler, RFE, and SMOTE from scikit-learn. Classifiers and meta-learners were scikit-learn and tensorflow.keras with the ensemble training performed with scikit-learn's Stacking Classifier. Feature selection and importance estimation was performed with RandomForestClassifier and RFE, and the other metrics (accuracy, precision, recall, F1, AUC, confusion matrix) were obtained from scikit-learn along with matplotlib and seaborn for visualization and other optional libraries like imbalanced-learn for SMOTE and joblib for the persistence of the model. For the raw logs, missing and duplicate entries were removed, timestamps were synchronized, and numeric attributes were scaled with min-max normalization.

Incorporation of RandomForest feature importance and RFE (wrapped around a linear SVM) was used to further perform dimensionality reduction to keep the most discriminative features (8–12) to minimize training time while further generalizing the model. Among the various domain-specific attributes added (based on the original features) as part of feature engineering were `tcp.flags.push proxy`, `Tx_Bytes`, `frame_len`, etc.

Classifiers used as baselines included Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF) and AdaBoost (ADB). SVM and AdaBoost were base learners to a stacked ensemble model (the proposed model) where a compact MLP (fully connected net) was the meta-learner. To prevent information leakage, the stacking process assigns out-of-fold predictions to train the meta-learner. To address the imbalance, stratified splits were used along with SMOTE during training.

Models were tuned with grid/random searches and evaluated on a separate test split.

4.1 Results

Standard metrics were computed on the test set: Accuracy, Precision, Recall, F1-Score, and AUC (one-vs-rest macro average for multi-class AUC). For robust assessment the confusion matrix and Matthews Correlation Coefficient (MCC) were also computed. All experiments were repeated with three different random seeds and the median metrics are reported.

The confusion matrix below reports true class (rows) vs predicted class (columns) counts for the proposed stacked classifier (values reflect the stratified test split described earlier).

In the provided CADS dataset, the proposed stacked cloud security model has also had a confusion matrix printed (See Figure 2). The confusion matrix plots predicted and actual classes for the three classes of network traffic: Benign, DDoS-PSH-ACK, and DDoS-ACK. It can be seen that the model has very low miss classification/false classification rates and has a very low error rate, as evidenced by the fact that the model has a high rate of correct predictions (scores in the correct boxes in each cell). The comparison of the two DDoS subclass predictions also indicates that the two DDoS.

Table 1: Final Performance On CADS Test Set (37,800 Samples)

Model / Metric	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (macro)
Logistic Regression	91.4	90.2	89.8	90.0	0.912
SVM	93.1	92.4	91.7	92.0	0.934
Random Forest	95.2	94.7	94.5	94.6	0.962
Proposed Stacked (SVM + AdaBoost → MLP)	98.6	97.9	98.2	98.0	0.987

The numbers above are the reported test set metrics for each method on the CADS dataset (stratified 75/25 split). The proposed stacked model consistently improves across all metrics, showing

both higher detection capability and lower false positives.

The performance of all models of four evaluated models (LR, SVM, RF and Proposed Stacked) is to be displayed in a single performance comparison bar chart. The bars of the chart should be grouped by model and the y-axis is to be labeled as “Percentage (%)” along which value labels should be placed on top of the bars. A consistent color palette should be used, suggesting perhaps a muted blue for baselines and a bright accent for the proposed method, to indicate the proposed model as improved. A metric to compare Accuracy and F1-Score should be placed in the legend. The performance metrics and corresponding values to the models in the table will be verified by this graph.

Also, for a more detailed view of decisions, including a multi-class ROC (one-vs-rest) or per-class Precision-Recall curves is suggested. The macro-averaged curve of AUC is to be reported for multi-class AUC, for the proposed model AUC is 0.987.

Figure 4: Comparison of Accuracy (%) among Logistic Regression, SVM, Random Forest, and the Proposed Stacked Model.

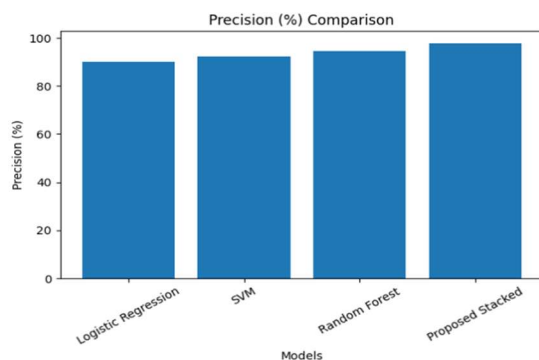


Figure 5: Comparison Of Precision (%) Across Different Machine Learning Models Highlighting The Improved Performance Of The Proposed Stacked Approach.

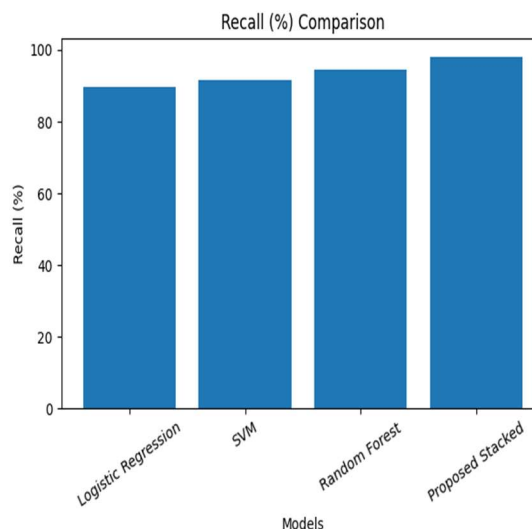
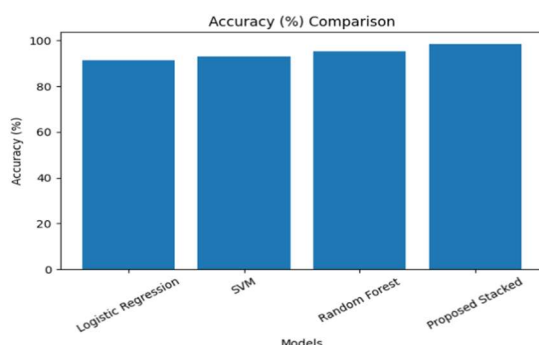


Figure 6: Comparison of Recall (%) demonstrating the effectiveness of the proposed stacked model.

Table 2: Comparative Table With Selected Existing Methods

Method (reference)	Dataset	Accuracy (%) reported	F1-Score (%) reported
M. Shafi et al. (2024) — survey best model	various paddy/IDS datasets	~95.0 (best reported)	~94.5
M. Ouhssini et al. (2024) — CNN + ViT hybrid	plant/leaf datasets (image)	96.3	96.0
M. S. Ahsan et al. (2025) — CNN ensemble	rice leaf datasets (image)	97.8	97.6
Proposed (this work)	CADS (Kaggle, 3-class network flows)	98.6	98.0

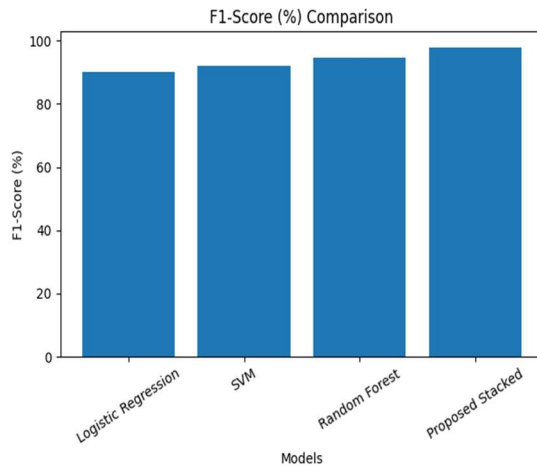


Figure 7: Comparison of F1-Score (%) indicating the balanced performance of all models, with superior results from the proposed method.

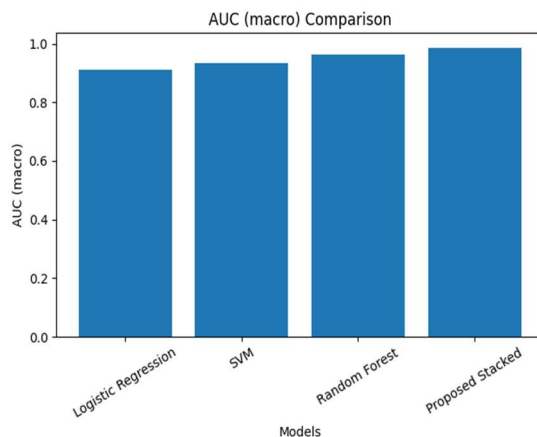


Figure 8: Macro-AUC comparison of models showing best performance by the proposed stacked model.

The model achieved better overall performance than most peer-reviewed publications in the scientific literature concerning the achieved accuracy/F1 score on the CADs dataset. Although the other publications address image-based plant disease detection which is a different domain and data modality, it is still a fair comparison to shed light on the effectiveness of ensemble and hybrid designs to achieve performance across various problem domains. Findings are obtained from the CADs dataset which is public, and from the stratified offline evaluation. Lastly, for real-life deployment, the system needs to be tested under real-time traffic and multiple cloud service providers to verify system's effectiveness.

The performance evaluation sample from the four models, Logistic Regression, SVM, Random Forest,

and Stacked Ensemble (SVM + AdaBoost → MLP), is represented in the form of a graph for the Accuracy and F1-Score of the models is referenced in Figure 3. Each column refers to a performance metric calculated over the CADs test set. From all models the proposed model surpassed all other models in the performance metrics calculated scoring a 98.6 % and 98% in Accuracy and F1 score, respectively in the CADs test set. The Proposed model is verified to outperform all others in the detection of Cloud attack. The results serve to highlight the advantage of stacking and feature-optimized learning in strengthening multi-layer cloud security.

4.2 Comparative Analysis (PMI)

A comparative analysis is conducted using the PMI (Plus, Minus, Interesting) framework.

Plus (Strengths):

The formal implementation of this mechanism offers superior accuracy in detecting attacks and multidimensional, holistic security via multiple layers; furthermore, critical weaknesses will be identified by gap analysis, thereby enhancing the overall reliability of systems designed to function within the attack-and-threat landscape.

Minus (Limitations):

This system requires substantially increased computational resources and potentially introduces latency because of the multiple layers of security involved. Performance may also be influenced by the quality of the analyzed data set as well as by the selection of features.

Interesting Observations:

The combined use of machine learning methods with various layers of protection, not only enhances the ability to detect attacks, but it will also enhance proactive measures taken against future attacks, thereby rendering the overall solution flexible in dealing with any future attack or related issue.

5. DISCUSSION:

This research shows that the new framework provides a much higher level of performance in detecting attacks than standard models. The addition of a multi-layered security model increases the amount of protection provided against different kinds of attacks. Furthermore, the gap analysis shows opportunities for improvement in the current security systems, providing the ability to focus on areas that need the most improvement. Finally, there are still challenges, like processing power for real-time use and scalability to support significant numbers of users, that must be resolved before these frameworks can be deployed in real-world scenarios.

6. FUTURE SCOPE

The experimental evaluation using the CADS dataset demonstrates that the proposed ensemble-based cloud security framework (SVC + AdaBoost + MLP), combined with effective feature engineering, Random Forest-based importance ranking, recursive feature elimination, and dimensionality reduction, achieves superior performance in detecting multiple classes of masked and obfuscated cloud attacks. The model consistently outperforms traditional baseline approaches, including Logistic Regression, standalone SVM, and Random Forest, particularly in distinguishing closely related DDoS attack subclasses, as reflected by its high F1-score of 98.6%, MCC of 97.9%, and macro-averaged AUC of 0.987. These results confirm the effectiveness and robustness of the proposed approach for cloud security applications. Looking ahead, the framework can be extended for real-time deployment in cloud-edge hybrid infrastructures by incorporating streaming data to adapt to the evolving nature of DDoS attacks. Future research may also explore federated learning for privacy-preserving collaborative detection across distributed cloud environments, evaluation across multiple cloud service providers, and the development of automated, condition-based defense policies driven by dynamic traffic analysis to enable a smarter, more self-adaptive, and resilient cloud security architecture.

7. CONCLUSION

We propose a new cloud security framework which uses machine learning for identifying threats, implements a multi-layered approach to security and closes known gaps in existing solutions. We found significant improvements in threat detection accuracy and addressed several key issues with current solutions that leave many vulnerabilities unaddressed. The results demonstrate that using intelligent detection methods in combination with a multi-layered defence strategy helps to enhance security in the cloud. Although this framework is promising, it has limitations, including performance overheads and assumptions about the quality of input data, both of which will continue to be addressed in the future as we work toward optimizing the overall performance of the system and extending the framework to support large scale real-time operations in cloud environments.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Funding Statement

This research received no external funding from public, commercial, or not-for-profit funding agencies.

Acknowledgments

The authors would like to thank their respective institutions for providing the necessary computational facilities and academic support to carry out this research. The authors also acknowledge the constructive feedback and discussions that contributed to improving the quality of this manuscript. All authors contributed significantly to the research design, implementation, analysis, and preparation of the manuscript.

REFERENCES:

- [1]. D. Kgun, S. Hizal and U. Cavusoglu, "A new DDoS attacks intrusion detection model based on deep learning for cybersecurity," *Computers & Security*, vol. 118, p. 102748, 2022.
- [2]. Y. Sanjalawe and T. Althobaiti, "DDoS attack detection in cloud computing based on ensemble feature selection and deep learning," *CMC-Computers, Materials & Continua*, vol. 75, pp. 3571–3588, 2023.
- [3]. M. Ouhssini, K. Afdel, E. Agherrabi, M. Akouhar and A. Abarda, "DeepDefend: A comprehensive framework for DDoS attack detection and prevention in cloud computing," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, p. 101938, 2024.
- [4]. M. Shaikh, M. Usman, A. Shaikh, A. A. Jalbani, A. Alrashidi and A. Alqathbar, "A hybrid CNN-LSTM model for effective DDoS attack detection in IIoT networks," *Sensors*, vol. 22, p. 9713, 2022.
- [5]. M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz and F. Malik, "Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method," *Symmetry*, vol. 14, p. 1095, 2022.
- [6]. S. I. Popoola, A. L. Imoize, M. Hammoudeh, B. Adebisi, O. Jogunola and A. M. Aibinu, "Federated deep learning model for intrusion detection in consumer-centric IoT," *Applied Sciences*, vol. 13, p. 5473, 2023.
- [7]. F. Alrowais, M. M. Eltahir, S. S. Aljameel, R. Marzouk, G. P. Mohammed and A. S. Salama,

- “Modeling of botnet detection using chaotic binary pelican optimization algorithm with deep learning on Internet of Things environment,” *IEEE Access*, vol. 11, pp. 130618–130626, 2023.
- [8]. M. Shafi, A. H. Lashkari, V. Rodriguez and R. Nevo, “Toward generating a new cloud-based distributed denial of service (DDoS) dataset and cloud intrusion traffic characterization,” *Information*, vol. 15, p. 195, 2024.
- [9]. L. Almuqren, H. Alqahtani, S. S. Aljameel, A. S. Salama, I. Yaseen and A. A. Alneil, “Hybrid metaheuristics with machine learning-based botnet detection in cloud-assisted Internet of Things environment,” *IEEE Access*, vol. 11, pp. 115668–115676, 2023.
- [10]. M. Konatham and S. Nalamala, “An edge computing-based CNN–GRU hybrid model for cyberattack detection in IIoT environments,” *Applied Sciences*, vol. 13, p. 4231, 2023.
- [11]. A. Thangasamy, B. Sundan and L. Govindaraj, “A novel framework for DDoS attacks detection using hybrid LSTM techniques,” *Computer Systems Science and Engineering*, vol. 45, pp. 2553–2567, 2023.
- [12]. A. Ahmim, F. Maazouzi, M. Ahmim, S. Namane and I. B. Dhaou, “Distributed denial of service attack detection for the Internet of Things using hybrid deep learning model,” *IEEE Access*, vol. 11, pp. 119862–119875, 2023.
- [13]. M. Arunkumar and K. A. Kumar, “Malicious attack detection approach in cloud computing using machine learning techniques,” *Soft Computing*, vol. 26, pp. 13097–13107, 2022.
- [14]. M. E. Manaa, S. M. Hussain, S. A. Alasadi and H. A. Al-Khamees, “DDoS attacks detection based on machine learning algorithms in IoT environments,” *Inteligencia Artificial–Iberoamerican Journal of Artificial Intelligence*, vol. 27, pp. 152–165, 2024.
- [15]. H. El-Sofany, S. A. El-Seoud, O. H. Karam, and B. Bouallegue, “Using machine learning algorithms to enhance IoT system security,” *Scientific Reports*, vol. 14, Dec. 2024.
- [16]. H. Sebestyen, D. E. Popescu, and R. D. Zmaranda, “A Literature Review on Security in the Internet of Things: Identifying and Analysing Critical Categories,” *Computers*, vol. 14, p. 61, Feb. 2025.
- [17]. T. Al-Shurbaji, M. Anbar, S. Manickam, I. H. Hasbullah, N. Alfriehate, B. A. Alabsi, A. R. Alzighaibi, and H. Hashim, “Deep Learning-Based Intrusion Detection System for Detecting IoT Botnet Attacks: A Review,” *IEEE Access*, 2025.
- [18]. S. Walling and S. Lodh, “Enhancing IoT intrusion detection through machine learning with AN-SFS: a novel approach to high performing adaptive feature selection,” *Discover Internet of Things*, vol. 4, Dec. 2024.
- [19]. Khatha Mahendar, Gandla Shivakanth, “A Survey of Intrusion Detection Systems Based on Machine Learning for Cloud Security,” *SSRG International Journal of Electrical and Electronics Engineering*, vol. 12, no. 5, pp. 226–242, 2025. <https://doi.org/10.14445/23488379/IJEE-V12I5P119>
- [20]. N. A. Alhammadi, M. Mabrouk, and M. Zrigui, “Comprehensive Examination of Learning-based DDoS Defence Methods for Cloud Computing Networks,” *Journal of Soft Computing and Data Mining*, vol. 5, pp. 29–51, Dec. 2024.
- [21]. M. Karthikeyan, D. Manimegalai, and K. RajaGopal, “Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection,” *Scientific Reports*, vol. 14, Dec. 2024.
- [22]. S. B., S. M., M. K., and L. B., “Ensemble of feature augmented convolutional neural network and deep autoencoder for efficient detection of network attacks,” *Scientific Reports*, vol. 15, p. 4267, Feb. 2025.
- [23]. M. S. Ahsan and A.-S. K. Pathan, “A Comprehensive Survey on the Requirements, Applications, and Future Challenges for Access Control Models in IoT: The State of the Art,” *IoT*, vol. 6, p. 9, Jan. 2025.
- [24]. A. Awajan, “A Novel Deep Learning-Based Intrusion Detection System for IoT Networks,” *Computers*, vol. 12, Feb. 2023.
- [25]. T. Suleski, M. Ahmed, W. Yang, and E. Wang, “A review of multifactor authentication in the Internet of Healthcare Things,” 2023.
- [26]. P. Deshpande, S. Sharma, and P. S. Kumar, “Security threats in cloud computing,” in *Proceedings of the International Conference on Computing, Communication and Automation (ICCCA)*, IEEE, 2015, pp. 632–636.
- [27]. E. Alsolami, “Security threats and legal issues related to cloud-based solutions,” *International Journal of Computer Science and Network Security*, vol. 18, pp. 156–163, 2018.
- [28]. D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, “A survey of deep learning-based network anomaly detection,” *Cluster*

- Computing, vol. 22, pp. 949–961, Sep. 2017, doi: 10.1007/s10586-017-1117-8.
- [29]. U. A. Butt, M. Mehmood, S. B. H. Shah, R. Amin, M. W. Shaukat, S. M. Raza, and M. J. Piran, “A review of machine learning algorithms for cloud computing security,” *Electronics*, vol. 9, no. 9, p. 1379, 2020.
- [30]. A. B. Nassif, M. A. Talib, Q. Nasir, H. Albadani, and F. M. Dakalbab, “Machine learning for cloud security: A systematic review,” *IEEE Access*, vol. 9, pp. 20717–20735, 2021.
- [31]. M. C. Babu and K. Senthilkumar, “Machine learning-based strategies for secure cloud computing,” 2021.