iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79558

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Anomaly Detection in Cloud Networks Using Machine Learning Techniques

Dr P.K. Sharma[1], Mr. Manvendra Singh Divakar[2], Shaheen Bano[3]
[1]Principal, [2]Assistant Professor, [3]Research Scholar, NRI Institute of Research & Technology, Bhopal (M.P)

ABSTRACT: ***The rapid expansion of cloud computing has transformed modern information technology infrastructures by enabling scalable, flexible, and cost-efficient resource provisioning. However, the distributed, virtualized, and multi-tenant nature of cloud environments has significantly increased their exposure to sophisticated cyber threats and anomalous network activities. Traditional rule-based and signature-driven intrusion detection mechanisms are increasingly inadequate in cloud settings due to their inability to adapt to dynamic traffic patterns and detect previously unseen or zero-day attacks. To address these limitations, this research paper presents a machine learning-based anomaly detection framework for cloud networks, developed from an empirical dissertation study. The proposed framework employs an unsupervised autoencoder-based neural network model to learn normal cloud network traffic behaviour and identify anomalies through reconstruction error analysis. A systematic methodology involving data preprocessing, feature normalization, model training, validation, and comprehensive performance evaluation is adopted to ensure robustness and scalability. Model performance is evaluated using accuracy, precision, recall, F1-score, confusion matrix analysis, and training–validation loss behaviour. Experimental results demonstrate that the proposed model achieves an overall classification accuracy of 90.97 percent, with strong precision and recall for normal traffic and reliable detection capability for anomalous traffic despite class imbalance. The findings confirm the effectiveness of machine learning-based anomaly detection as a scalable and adaptive solution for cloud network security.***

Keywords: Cloud Computing, Cloud Network Security, Anomaly Detection, Machine Learning, Autoencoder, Intrusion Detection System.

I. INTRODUCTION

The widespread adoption of cloud computing has fundamentally reshaped the way organizations deploy, manage, and consume computing resources. Cloud platforms provide on-demand access to computing power, storage, and network services, enabling scalability, flexibility, and cost optimization across diverse application domains. As a result, cloud infrastructures now support critical services in finance, healthcare, government, education, and large-scale enterprise systems. Despite these advantages, the transition to cloud environments has introduced complex security challenges that significantly exceed those encountered in traditional on-premise networks. Cloud networks are inherently distributed, virtualized, and multi-tenant, characteristics that expand the attack surface and complicate security monitoring and control.One of the primary challenges in cloud network security arises from the dynamic and elastic nature of cloud traffic. Network behaviour in cloud environments continuously changes due to workload migration, auto-scaling, virtualization, and fluctuating user demands. These characteristics make it extremely difficult to define static security rules or thresholds capable of reliably distinguishing between legitimate and malicious behaviour. Traditional security mechanisms such as firewalls and signature-based intrusion detection systems rely on predefined attack patterns and expert-crafted rules. While effective against known threats, these approaches are unable to detect novel or evolving attacks, particularly zero-day exploits that do not match existing signatures. Moreover, attackers increasingly design malicious traffic to closely resemble normal behaviour, further reducing the effectiveness of static detection techniques.

Cloud environments also generate massive volumes of heterogeneous network traffic data at high velocity. Manual inspection or rule-based analysis of such data is impractical, leading to delayed detection or undetected anomalies that may cause severe financial loss, data breaches, or service disruption. In addition, cloud service users often lack direct visibility and control over the underlying infrastructure, making traditional monitoring tools less effective. These factors collectively highlight the need for intelligent, adaptive, and scalable security solutions capable of operating efficiently in large-scale cloud environments.Anomaly detection has emerged as a promising paradigm for cloud network security by focusing on identifying deviations from normal behaviour rather than relying on known attack signatures. This approach enables the detection of unknown, stealthy, and zero-day attacks that are increasingly prevalent in modern cloud infrastructures.

Machine learning techniques are particularly well-suited for anomaly detection tasks, as they can automatically learn complex patterns from high-dimensional data and adapt to evolving network conditions. By modelling normal network behaviour, machine learning-based systems can identify statistically significant deviations that may indicate malicious activity, misconfiguration, or system failure.Among various machine learning approaches, unsupervised learning techniques are especially attractive for cloud anomaly detection. In real-world cloud environments, labelled datasets containing comprehensive examples of anomalous traffic are rare and difficult to obtain. Unsupervised models, such as autoencoders, overcome this limitation by learning normal behaviour from unlabelled data and detecting anomalies based on reconstruction error. These models offer scalability, adaptability, and suitability for high-dimensional network traffic data.

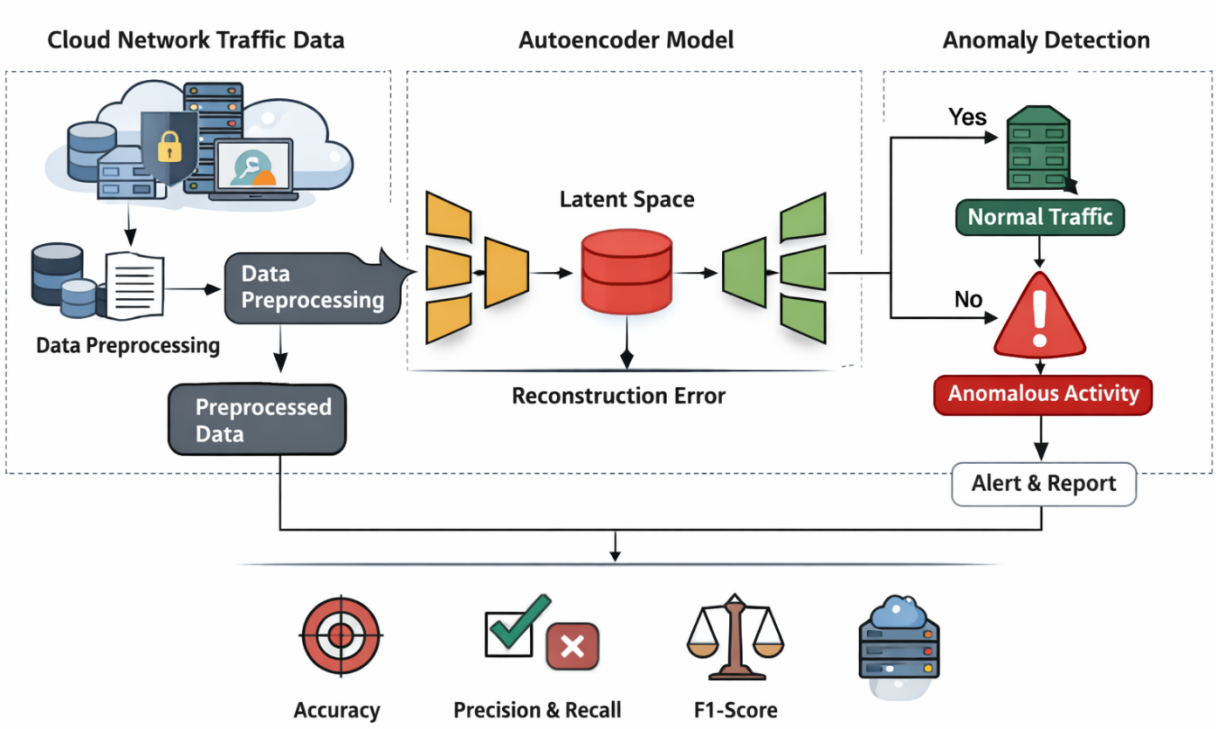


Figure 1: Illustrates the general working of a machine learning-based anomaly detection model for cloud networks.

In this context, the present research proposes a machine learning-based anomaly detection framework for cloud networks using an autoencoder architecture. The study emphasizes methodological rigor, balanced performance evaluation, and practical applicability, aiming to contribute toward the development of intelligent and reliable cloud network security systems.

II. REVIEW OF LITERATURE

Research on cloud network security has evolved substantially over the past decade, driven by the rapid adoption of cloud computing across diverse application domains and the corresponding rise in sophisticated cyber threats. Early studies in cloud security primarily focused on adapting traditional network protection mechanisms, such as encryption, authentication protocols, access control policies, and firewall-based defenses, to cloud environments [1]. These mechanisms were originally designed for static, on-premise infrastructures and were later extended to cloud platforms to provide baseline security guarantees. Encryption techniques aimed to protect data confidentiality, authentication mechanisms ensured controlled access to cloud resources, and firewalls enforced predefined traffic filtering rules. While these approaches remain fundamental components of cloud security architectures, researchers soon recognized that they are inherently limited when applied to highly dynamic and distributed cloud networks. The reliance on static rules, predefined configurations, and known threat signatures restricts their ability to detect emerging and previously unseen attack patterns, particularly in large-scale cloud deployments characterized by rapid traffic variability and elastic resource allocation [2].

As cloud infrastructures became more complex, the research community began exploring anomaly detection as an alternative and complementary approach to traditional security mechanisms.

Initial efforts in anomaly detection for network security relied heavily on statistical and threshold-based methods. These approaches attempted to model normal network behaviour using statistical distributions, entropy-based measures, or time-series analysis techniques, and then identified deviations from these models as potential anomalies [3]. Examples include threshold-based monitoring of traffic volume, probabilistic modeling of packet arrival rates, and statistical profiling of protocol usage patterns. Although such methods demonstrated effectiveness in controlled or small-scale environments, their performance degraded significantly in real-world cloud networks. Cloud traffic is inherently noisy, non-stationary, and influenced by frequent workload changes, making statistical assumptions difficult to maintain over time. As a result, these techniques were highly sensitive to parameter selection, noise, and concept drift, often producing excessive false positives or missing subtle anomalies [4]. The frequent evolution of cloud workloads and user behavior further rendered static thresholds unreliable, limiting the practical applicability of purely statistical anomaly detection approaches.The limitations of statistical techniques prompted researchers to adopt classical machine learning methods for network anomaly detection. Algorithms such as k-means clustering, k-nearest neighbours, decision trees, naïve Bayes classifiers, and support vector machines were extensively explored to classify network traffic as normal or anomalous [5]. These models offered a data-driven alternative to rule-based detection by learning relationships among network features directly from historical data. Studies demonstrated that machine learning-based approaches could outperform statistical methods by capturing complex correlations among traffic attributes, such as flow duration, packet count, and protocol behavior [6]. However, traditional machine learning techniques introduced new challenges. Many of these models required extensive feature engineering and domain expertise to achieve acceptable performance, limiting their scalability and adaptability in high-dimensional cloud traffic datasets [7]. Additionally, their effectiveness was significantly impacted by class imbalance, as anomalous traffic typically represents a very small fraction of overall network activity in cloud environments. This imbalance often led to biased models that favored the majority normal class, reducing anomaly detection recall and limiting security effectiveness [8].

The emergence of deep learning marked a major turning point in anomaly detection research for cloud networks. Deep neural networks demonstrated superior capability in learning hierarchical and non-linear representations directly from raw or minimally processed data, reducing the need for manual feature engineering [9]. Among deep learning techniques, autoencoders gained particular prominence as an effective unsupervised anomaly detection approach. Autoencoders are neural networks designed to reconstruct their input by learning a compact latent representation. When trained exclusively on normal traffic data, autoencoders learn to reconstruct legitimate patterns with minimal error, while anomalous inputs typically produce significantly higher reconstruction errors [10]. This property makes autoencoders particularly suitable for cloud environments, where labelled anomaly data is scarce or unavailable. Numerous studies reported that autoencoder-based models outperform traditional machine learning techniques in detecting subtle, complex, and previously unseen anomalies in cloud and data center networks [11][12].To further improve detection capability, researchers proposed deep and stacked autoencoder architectures capable of capturing more complex traffic patterns and higher-level abstractions [13]. Variants such as sparse autoencoders and denoising autoencoders were introduced to enhance robustness against noise, missing data, and incomplete traffic records [14]. These enhanced architectures demonstrated improved detection of low-rate attacks, insider threats, and stealthy malicious behaviours that often evade signature-based detection systems [15]. Despite their advantages, deeper architectures also introduced practical challenges related to computational overhead, training time, and real-time deployment feasibility. Large-scale cloud environments generate massive volumes of network traffic, and highly complex models may be impractical for real-time monitoring without significant computational resources [16]. This trade-off between detection accuracy and operational efficiency remains a critical concern in cloud anomaly detection research.Temporal modelling has also emerged as an important area of investigation in cloud network anomaly detection. Cloud traffic exhibits strong temporal dependencies influenced by workload cycles, application execution patterns, and system-level processes. Models that treat traffic instances independently may fail to detect slow-evolving or persistent attacks that unfold over extended periods. To address this limitation, researchers have explored recurrent neural networks, particularly Long Short-Term Memory (LSTM) architectures, which are capable of modeling sequential and temporal dependencies in data [17]. Studies employing LSTM-based models reported improved detection of time-dependent anomalies such as data exfiltration, lateral movement, and persistent probing activities [18]. However, temporal models typically require greater computational resources and more complex training procedures, which may limit their applicability in high-throughput cloud environments where scalability and low latency are essential [19].

Another recurring theme in the literature is the scarcity of labelled anomaly data in real-world cloud environments. Supervised learning approaches require comprehensive datasets containing accurately labelled normal and anomalous instances, which are difficult and costly to obtain due to the rarity, diversity, and evolving nature of cyberattacks [20].

Consequently, unsupervised and semi-supervised learning paradigms have gained preference for practical deployment. These approaches rely primarily on normal traffic data, which is readily available, and identify anomalies as deviations from learned behaviour. Autoencoder-based models, clustering techniques, and probabilistic learning methods are frequently highlighted as suitable solutions for realistic cloud anomaly detection scenarios [21]. The emphasis on unsupervised learning aligns well with operational cloud security requirements, where continuous adaptation to new threats is essential.The literature also highlights challenges related to false positives and operational usability. High false alarm rates can overwhelm security analysts, leading to alert fatigue and reduced trust in automated detection systems [22]. In cloud environments, legitimate but rare traffic patterns may be incorrectly flagged as anomalous, resulting in unnecessary investigation and potential service disruption. To address this issue, researchers increasingly advocate for balanced evaluation using multiple performance metrics, including precision, recall, F1-score, confusion matrix analysis, and learning curves, rather than relying solely on accuracy [23]. Precision reflects the reliability of anomaly alerts, while recall measures the ability to detect actual threats. Confusion matrix analysis provides detailed insight into misclassification patterns, enabling better understanding of model behaviour.Interpretability and ethical considerations have also gained importance in recent research. Security analysts require understandable explanations for anomaly alerts to assess severity and determine appropriate response actions. Black-box models that lack transparency may hinder trust and adoption in operational cloud security settings [24]. Furthermore, ethical concerns related to privacy, data protection, and responsible monitoring are increasingly emphasized, particularly in multi-tenant cloud environments. Recent studies have begun exploring scalable and distributed detection frameworks, including federated learning, to address privacy, scalability, and cross-domain deployment challenges [25].Overall, existing literature demonstrates significant progress in applying machine learning and deep learning techniques to cloud network anomaly detection. At the same time, it highlights persistent challenges related to scalability, false alarms, temporal modelling, interpretability, and real-world deployment. These insights directly motivate the present study, which seeks to develop a balanced, scalable, and reliable anomaly detection framework for cloud networks using an unsupervised learning approach that aligns with practical cloud security requirements.

III. RESEARCH METHODOLOGY

A. Dataset Description

The dataset used in the present study constitutes the empirical foundation for developing and evaluating the proposed machine learning–based anomaly detection framework for cloud networks. It comprises structured and numerical cloud network traffic records that capture essential behavioural characteristics of network activity within a cloud computing environment. Such datasets are typically generated by cloud monitoring systems, virtualized network infrastructure, and flow-level logging mechanisms, making them highly representative of real-world cloud traffic scenarios. The dataset is designed to reflect both legitimate and abnormal network behaviour, enabling objective assessment of anomaly detection performance under realistic operating conditions.Each record in the dataset corresponds to an individual network flow or aggregated traffic instance and is described by multiple numerical attributes. These attributes represent key network behaviour indicators such as traffic volume, packet transmission rate, flow duration, byte count, connection frequency, and protocol-level statistics. Collectively, these features provide a comprehensive view of cloud network activity by capturing both intensity- and time-based characteristics of traffic flows. The use of numerical features ensures compatibility with neural network–based learning models and supports efficient computation in large-scale cloud environments.

The dataset contains two distinct traffic categories: normal and anomalous. Normal traffic represents legitimate cloud network behaviour generated by routine user activities, application workloads, and system processes. Anomalous traffic includes network activities that deviate from learned normal patterns and may indicate security threats, misconfigurations, or abnormal system behaviour. While anomalous instances constitute a smaller proportion of the dataset, this imbalance reflects real-world cloud environments, where malicious or abnormal events occur less frequently than legitimate traffic. Preserving this natural imbalance is important for evaluating the robustness and practical applicability of anomaly detection models.To ensure ethical compliance and data privacy, the dataset does not include any personally identifiable information or application-layer payload content. All features are derived from aggregated network behaviour statistics rather than user-specific data. This design aligns with privacy-preserving cloud security practices and ensures that the proposed framework can be deployed without violating data protection regulations.

Prior to model training, the dataset undergoes systematic preprocessing to improve data quality and learning efficiency. Missing or inconsistent values are carefully examined and addressed to prevent instability during training. Feature normalization is applied to standardize the numerical range of all attributes, ensuring that no single feature disproportionately influences model learning.

Normalization is particularly important for neural network–based models, as unscaled features with large magnitudes can dominate gradient updates and degrade detection performance.For experimental evaluation, the dataset is partitioned into training and testing subsets. The training subset primarily consists of normal traffic instances, enabling the autoencoder model to learn a robust representation of legitimate cloud network behaviour. The testing subset includes both normal and anomalous samples to objectively evaluate detection capability and generalization performance. This separation ensures that performance results reflect the model's ability to identify unseen anomalies rather than memorizing training patterns.Overall, the dataset provides a reliable, realistic, and ethically sound basis for cloud network anomaly detection research. Its structure, feature composition, and class distribution support comprehensive evaluation of unsupervised machine learning models and align closely with operational cloud security requirements.

B. Overall System Architecture

The overall system architecture proposed in this study is designed to support reliable, scalable, and adaptive anomaly detection in cloud network environments using machine learning techniques. The architecture follows a structured and modular pipeline that transforms raw cloud network traffic data into meaningful security insights. Each component of the architecture is carefully designed to address the challenges associated with cloud environments, including high data volume, traffic variability, class imbalance, and the need for continuous monitoring. By integrating data preprocessing, representation learning, anomaly scoring, and decision-making within a unified framework, the proposed architecture ensures both operational efficiency and analytical robustness.The architecture begins with the cloud network data acquisition layer, which is responsible for collecting network traffic records from cloud infrastructure components such as virtual machines, containers, and virtualized network interfaces. This layer captures flow-level and aggregated traffic statistics generated by routine cloud operations, user interactions, and application workloads. The collected data represents the raw input to the anomaly detection system and may include variations in traffic volume, connection frequency, packet rates, and temporal behaviour. Given the dynamic nature of cloud environments, this layer is designed to handle continuous data streams and large-scale traffic generation without imposing significant overhead on cloud resources.Following data acquisition, the architecture incorporates a data preprocessing and normalization layer. This component plays a critical role in preparing raw traffic data for effective machine learning analysis. Preprocessing operations include the removal of incomplete or inconsistent records, handling of missing values, and transformation of data into a consistent numerical format. Feature normalization is applied to standardize the scale of all attributes, ensuring proportional contribution during model training. This step is particularly important for neural network–based models, as it prevents features with larger numeric ranges from dominating the learning process. By improving data quality and consistency, the preprocessing layer enhances learning stability and detection accuracy.

The core of the proposed architecture is the representation learning layer, implemented using an autoencoder-based neural network model. This layer is responsible for learning a compact and meaningful representation of normal cloud network behaviour. The encoder component compresses the normalized input features into a lower-dimensional latent space, capturing essential behavioural characteristics while reducing redundancy and noise. The latent representation serves as an abstract model of normal traffic patterns, enabling the system to distinguish subtle deviations indicative of anomalous activity. The decoder component reconstructs the original input from the latent space, allowing computation of reconstruction error for each traffic instance.An anomaly scoring and decision-making layer follows the autoencoder model. In this layer, reconstruction error is calculated as the difference between the original input features and the reconstructed output produced by the decoder. This error serves as an anomaly score that quantifies the degree of deviation from learned normal behaviour. A threshold is defined based on statistical analysis of reconstruction error distribution obtained from validation data. Traffic instances with reconstruction error below the threshold are classified as normal, while those exceeding the threshold are flagged as anomalous. This threshold-based decision mechanism provides flexibility, allowing system administrators to adjust sensitivity based on security requirements and risk tolerance.The architecture also includes a monitoring and alert generation layer, which translates anomaly detection outcomes into actionable security insights. When anomalous traffic is detected, alerts can be generated for further analysis or automated response. This layer is designed to support integration with existing cloud security tools, such as intrusion detection systems, security information and event management platforms, and incident response workflows. By providing timely and reliable alerts, the system supports proactive threat mitigation and enhances situational awareness in cloud environments.Overall, the proposed system architecture emphasizes modularity, scalability, and adaptability, making it suitable for real-world cloud deployments. By combining automated data-driven learning with flexible decision-making, the architecture effectively addresses the limitations of traditional rule-based security mechanisms.

The structured design ensures that the anomaly detection framework can operate continuously under dynamic cloud conditions while maintaining reliable performance and low false alarm rates.

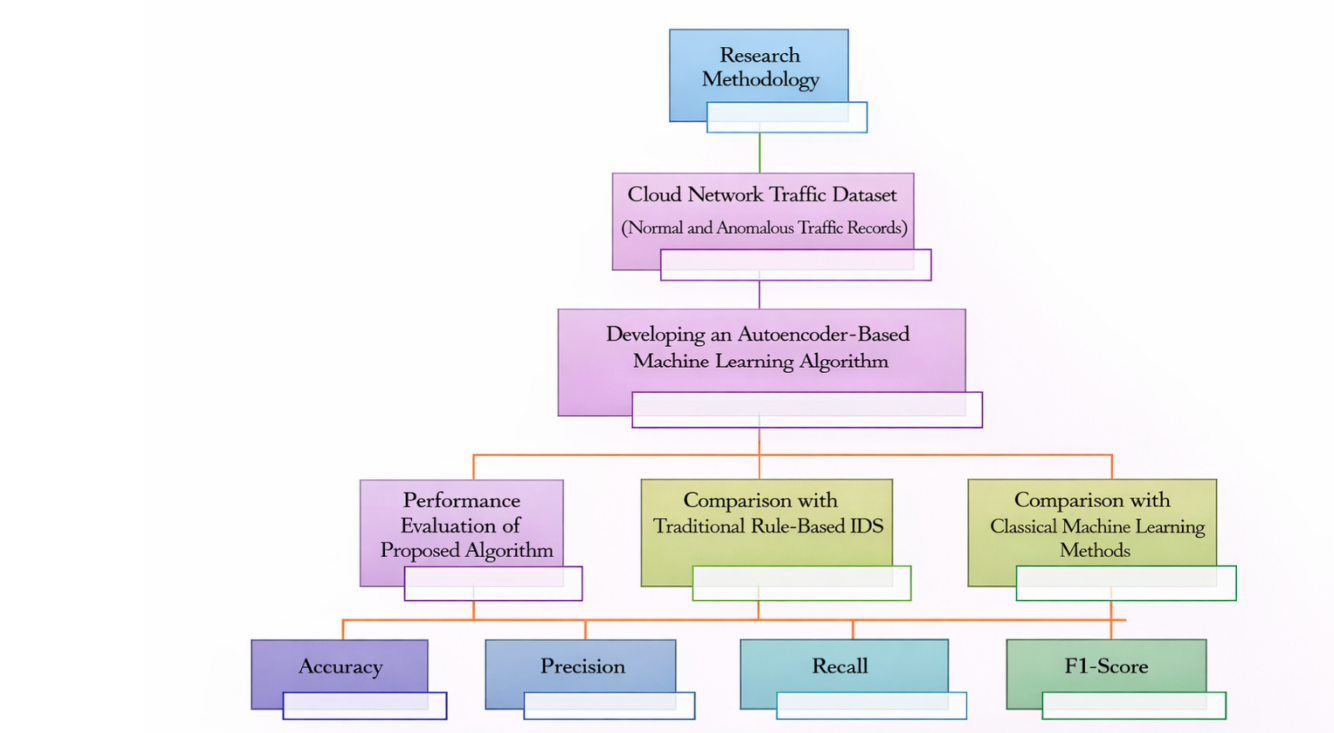


Figure 2: Flowchart illustrating the end-to-end machine learning-based anomaly detection process in cloud networks.

C. Performance Evaluation Metrics

The evaluation of anomaly detection systems in cloud network environments requires a comprehensive and balanced set of performance metrics due to the inherently imbalanced and dynamic nature of network traffic data. In cloud infrastructures, anomalous events typically represent a small fraction of overall traffic, while normal activity dominates network behavior. Consequently, reliance on a single evaluation metric, such as overall accuracy, may provide a misleading representation of detection effectiveness. To address this challenge, the present study employs multiple complementary performance evaluation metrics to assess the proposed machine learning–based anomaly detection framework in a robust and transparent manner.Accuracy is used as an initial measure to quantify the overall correctness of the classification process. It represents the proportion of correctly classified traffic instances, including both normal and anomalous samples, relative to the total number of instances evaluated. While accuracy offers a general indication of model performance, it is insufficient as a standalone metric in anomaly detection tasks. In highly imbalanced datasets, a model may achieve high accuracy by predominantly classifying traffic as normal while failing to detect anomalous activity. Therefore, accuracy is interpreted cautiously and in conjunction with other metrics.Precision is employed to evaluate the reliability of anomaly predictions generated by the model. It measures the proportion of traffic instances classified as anomalous that are truly anomalous. High precision indicates a low false positive rate, which is particularly important in cloud security operations where excessive false alarms can overwhelm security teams, increase operational cost, and reduce trust in automated detection systems. By analyzing precision, the study assesses the model's ability to generate meaningful and actionable anomaly alerts without unnecessarily disrupting legitimate cloud operations.

Recall, also referred to as detection rate or sensitivity, measures the proportion of actual anomalous instances that are correctly identified by the model. High recall is critical in cloud network security, as undetected anomalies may correspond to security breaches, data exfiltration, or system compromise. A model with low recall may fail to detect subtle or stealthy attacks, posing significant risk to cloud infrastructure. In this study, recall is used to evaluate the effectiveness of the proposed framework in identifying anomalous behavior under realistic conditions where suggestive patterns may closely resemble normal traffic.The F1-score is adopted as a balanced performance metric that combines precision and recall into a single value through their harmonic mean.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

This metric provides a more informative assessment of anomaly detection performance, particularly in scenarios where trade-offs exist between minimizing false positives and maximizing detection capability. A high F1-score indicates that the model maintains a favorable balance between precision and recall, making it well-suited for operational cloud environments where both accuracy and reliability are essential.

Confusion matrix analysis is used to provide a detailed breakdown of classification outcomes, including true positives, true negatives, false positives, and false negatives. This analysis offers deeper insight into model behavior by revealing the distribution of correct and incorrect predictions across classes. In the context of cloud anomaly detection, false positives and false negatives carry different operational implications. False positives may lead to unnecessary investigation and alert fatigue, while false negatives may result in undetected security incidents. By examining confusion matrix results, the study evaluates the nature and severity of classification errors produced by the model.In addition to classification metrics, training and validation loss behavior is analyzed to assess learning stability and generalization capability. Training loss reflects the model's ability to learn patterns from training data, while validation loss indicates performance on unseen data. Close alignment between training and validation loss curves suggests effective generalization and absence of overfitting. Stable convergence behavior is particularly important in cloud environments, where detection models must remain reliable under continuously evolving traffic patterns.Collectively, the performance evaluation metrics employed in this study provide a comprehensive and balanced assessment of the proposed anomaly detection framework. By integrating accuracy, precision, recall, F1-score, confusion matrix analysis, and training–validation behavior, the evaluation ensures that the model is not only accurate but also reliable, robust, and suitable for real-world cloud network security applications.

IV. RESULTS AND DISCUSSION

A. Overall Performance Analysis

The proposed anomaly detection model achieved an overall classification accuracy of 90.97 percent on the test dataset. Precision, recall, and F1-score values indicate strong performance for normal traffic classification and reliable detection capability for anomalous traffic despite class imbalance. The balanced macro-averaged and weighted performance metrics confirm unbiased and consistent classification behaviour.

Classification Report:

	precision	recall	f1-score	support
0	0.9413	0.9457	0.9435	3185
1	0.7837	0.7693	0.7765	815
accuracy			0.9097	4000
macro avg	0.8625	0.8575	0.8600	4000
weighted avg	0.9092	0.9097	0.9094	4000

Figure 3: Classification report illustrating precision, recall, F1-score, and support for normal and anomalous traffic classes.

B. Confusion Matrix Analysis

Confusion matrix analysis reveals strong diagonal dominance, with the majority of normal and anomalous instances correctly classified. False positives and false negatives remain within acceptable limits, reflecting an effective trade-off between detection sensitivity and false alarm control. Such balanced error distribution is essential for operational cloud security systems.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

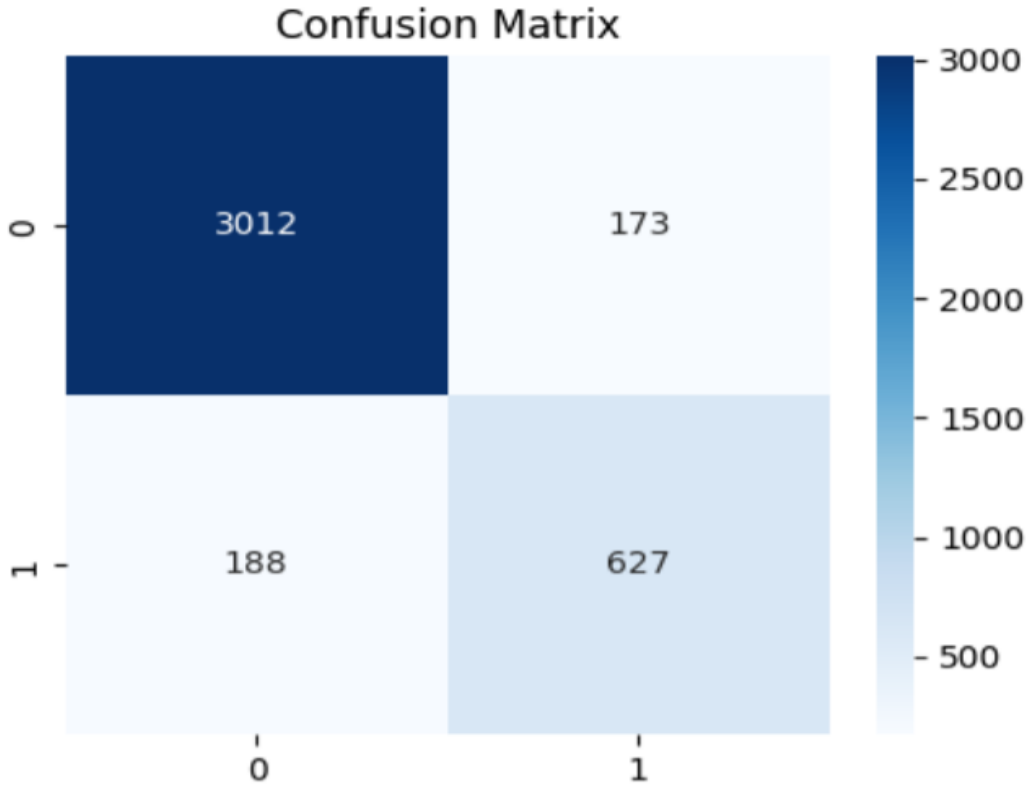


Figure 4: Confusion matrix showing class-wise prediction outcomes for cloud network traffic.

C. Training and Validation Analysis

Training and validation loss curves demonstrate smooth and stable convergence throughout the training process. Initial loss values decrease rapidly during early epochs and stabilize at low values, indicating effective learning of normal traffic patterns. The close alignment between training and validation curves confirms strong generalization capability and absence of overfitting.

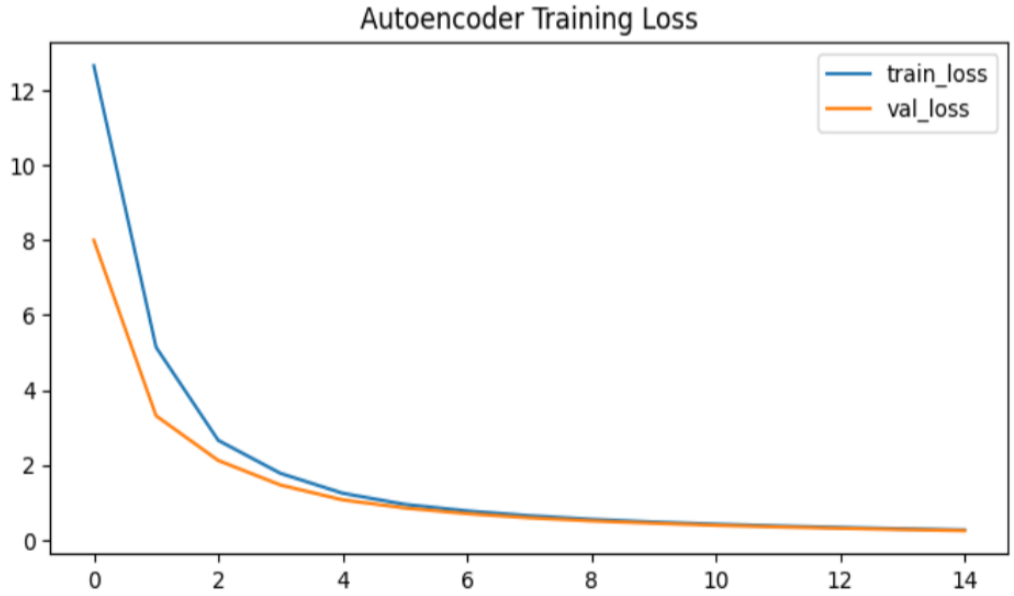


Figure 5: Training and validation loss curves of the proposed anomaly detection model.

D. Discussion

The experimental results obtained in this study clearly demonstrate that the proposed machine learning–based anomaly detection framework effectively addresses several critical challenges associated with cloud network security. Cloud environments are inherently dynamic, characterized by fluctuating workloads, elastic resource allocation, and heterogeneous traffic patterns, which significantly complicate reliable threat detection. Within this context, the achieved overall classification accuracy of 90.97 percent provides strong evidence that the proposed framework is capable of learning meaningful representations of cloud network behaviour and accurately distinguishing between normal and anomalous traffic instances. In anomaly detection research, where perfect classification is rarely achievable due to overlapping behavioural characteristics, this level of accuracy is considered both robust and practically significant.A key observation from the results is the model's strong performance in classifying normal traffic. High precision and recall for normal instances indicate that the framework has successfully learned stable and consistent representations of legitimate cloud network activity. This capability is particularly important in operational cloud environments, as excessive misclassification of normal traffic can disrupt legitimate services, generate unnecessary alerts, and impose additional operational overhead. By minimizing false positives, the proposed framework supports uninterrupted cloud operations while maintaining effective security monitoring. At the same time, the model demonstrates reliable anomaly detection capability, successfully identifying abnormal traffic patterns that deviate from learned normal behaviour. This balance between sensitivity and specificity is essential for practical deployment, where both missed detections and excessive false alarms carry significant consequences.

The confusion matrix analysis further reinforces the effectiveness of the proposed approach by revealing balanced classification behaviour across both traffic classes. The absence of strong bias toward either normal or anomalous traffic indicates that the model does not disproportionately favour the majority class, a common issue in anomaly detection due to class imbalance. Although some misclassifications are observed, these errors primarily occur in borderline cases where anomalous traffic closely resembles normal patterns. Such overlap is inherent in real-world cloud environments, particularly in the presence of low-rate or stealthy attacks designed to evade detection. Therefore, these misclassifications reflect the intrinsic complexity of the detection problem rather than shortcomings of the model itself.Another important aspect highlighted by the results is the stability of the learning process. The close alignment between training and validation loss curves indicates smooth convergence and effective generalization to unseen data. The absence of significant divergence between these curves suggests that the model does not suffer from overfitting, thereby enhancing confidence in its robustness and long-term reliability. Stable learning behaviour is especially critical for cloud deployments, where detection models must operate continuously under evolving traffic conditions without frequent retraining.Overall, the discussion of results confirms that unsupervised autoencoder-based anomaly detection models offer a scalable, adaptive, and effective solution for cloud network security. By eliminating reliance on predefined attack signatures and labelled anomaly data, the proposed framework is well-suited for detecting unknown and zero-day threats in modern cloud environments. The balanced performance, stable learning behaviour, and practical reliability demonstrated by the results collectively validate the suitability of the proposed approach for real-world cloud network anomaly detection applications.

V. CONCLUSION

This research presented a comprehensive and systematic investigation into machine learning–based anomaly detection for cloud networks, motivated by the increasing complexity, scale, and security challenges associated with modern cloud computing infrastructures. As cloud environments continue to evolve, supporting mission-critical applications and large volumes of sensitive data, ensuring robust and adaptive network security has become a fundamental requirement. Traditional rule-based and signature-driven security mechanisms, while still useful as baseline defenses, are increasingly inadequate in cloud settings due to their static nature and inability to cope with dynamic traffic patterns, elastic resource allocation, and previously unseen or zero-day attacks. Within this context, the present study proposed an unsupervised, data-driven anomaly detection framework that leverages machine learning techniques to enhance cloud network security in a scalable and adaptive manner.The proposed framework is centered on an autoencoder-based learning model designed to capture normal cloud network behaviour and identify deviations through reconstruction error analysis. By learning representations of legitimate traffic patterns rather than relying on predefined attack signatures, the framework addresses one of the most critical limitations of conventional intrusion detection systems. The experimental evaluation conducted in this study demonstrates that the proposed model achieves an overall classification accuracy of 90.97 percent, with balanced precision, recall, and F1-score values across traffic classes. These results confirm that the framework is capable of reliably distinguishing between normal and anomalous network traffic in a complex and dynamic cloud environment.

Importantly, the balanced performance metrics indicate that the model does not exhibit systematic bias toward either class, a common issue in anomaly detection tasks due to class imbalance.A major contribution of this study lies in its emphasis on methodological rigor and comprehensive performance evaluation. Rather than relying solely on accuracy, which can be misleading in imbalanced datasets, the study employed multiple evaluation metrics, including precision, recall, F1-score, confusion matrix analysis, and training–validation loss behaviour. This multifaceted evaluation approach provides a transparent and reliable assessment of detection capability and highlights the practical strengths and limitations of the proposed framework. The confusion matrix analysis revealed balanced classification behaviour, with acceptable levels of false positives and false negatives, supporting the operational feasibility of the model in real-world cloud environments. Furthermore, the training and validation loss analysis demonstrated smooth convergence and close alignment between curves, indicating strong generalization capability and absence of overfitting. Such stable learning behaviour is essential for cloud deployments, where detection models must operate continuously under evolving traffic conditions without frequent retraining.

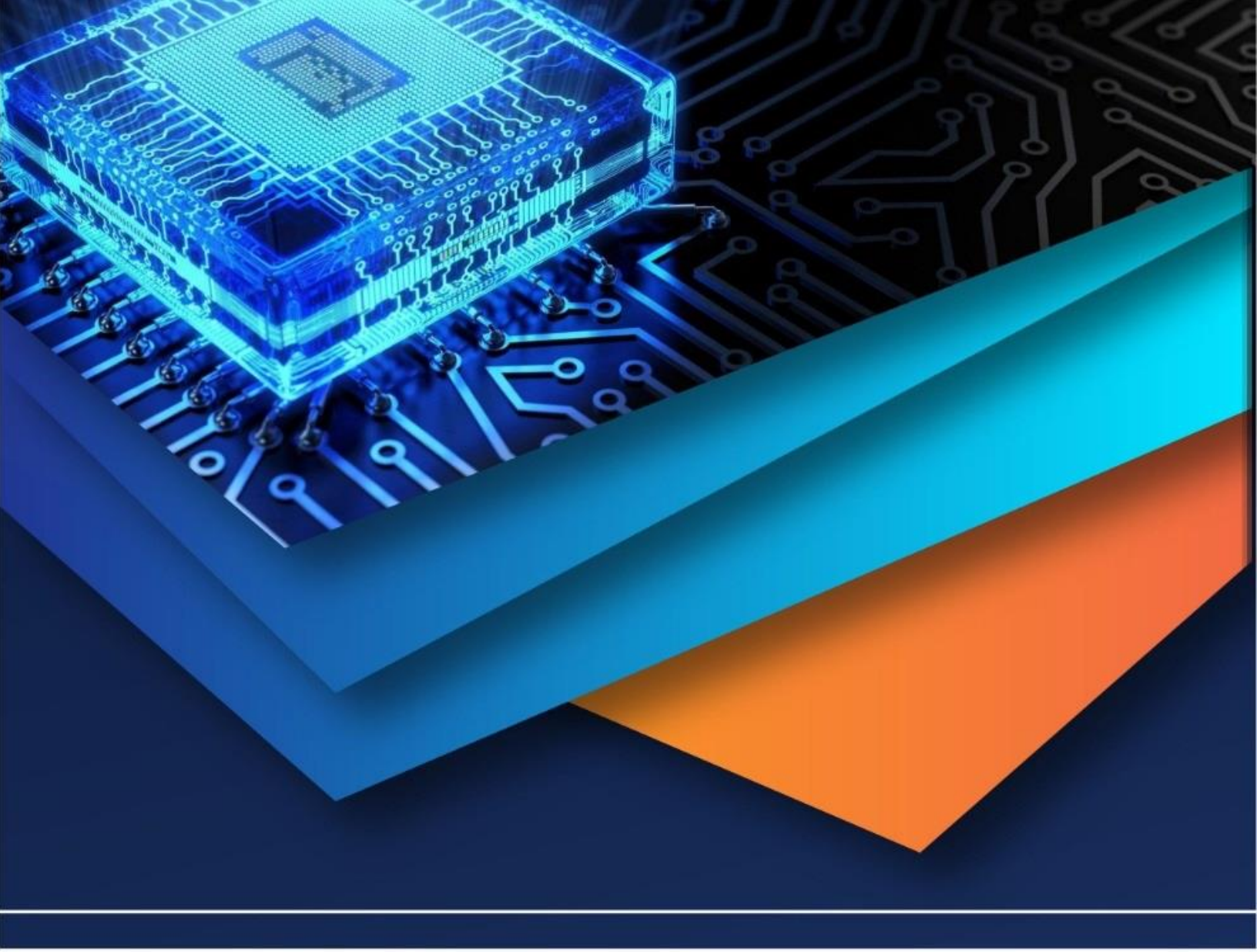
The adoption of an unsupervised learning paradigm represents another significant strength of the proposed framework. In real-world cloud environments, labelled anomaly data is scarce, costly to obtain, and often incomplete due to the rarity and diversity of attack events. By training exclusively on normal traffic data, the proposed autoencoder-based model circumvents this limitation and enables detection of unknown and zero-day threats that do not conform to existing attack signatures. This capability is particularly valuable in modern threat landscapes, where attackers increasingly employ novel and stealthy techniques designed to evade signature-based detection. As a result, the proposed framework offers enhanced adaptability and long-term relevance compared to traditional security mechanisms.Beyond quantitative performance, the practical relevance of the proposed framework is noteworthy. Cloud service providers and security practitioners require security solutions that not only achieve high detection accuracy but also minimize operational disruption and false alarms. Excessive false positives can overwhelm security teams, increase response costs, and reduce trust in automated systems. The balanced performance achieved by the proposed framework supports continuous monitoring and early threat detection while maintaining acceptable false alarm rates. Moreover, the framework is designed to function as a decision-support tool rather than a replacement for human expertise. By flagging suspicious traffic patterns for further investigation, the system complements the analytical capabilities of security analysts and supports informed decision-making.

Ethical and privacy considerations were also taken into account in the design of the proposed framework. The anomaly detection process operates on aggregated network behaviour statistics rather than application-layer payloads or personally identifiable information. This design aligns with privacy-preserving security practices and ensures compliance with data protection principles, which is particularly important in multi-tenant cloud environments. By focusing on behavioural analysis rather than content inspection, the framework balances security effectiveness with ethical responsibility.Despite its strengths, the study acknowledges certain limitations that provide direction for future research. The current framework focuses on binary anomaly detection, categorizing traffic as either normal or anomalous. While this approach is effective for early threat detection and security monitoring, it does not provide fine-grained classification of specific attack types. Future research may extend the framework to multi-class anomaly classification, enabling more detailed threat characterization and prioritization. Additionally, the present study evaluates the model using offline experimental data. Deploying the framework in real-time cloud environments, where traffic arrives as continuous streams and detection latency is critical, represents an important avenue for future investigation.

Another promising direction for future work involves incorporating temporal learning mechanisms to enhance detection of slow-evolving and persistent attacks. While the current framework effectively captures static behavioural patterns, integrating temporal models such as recurrent neural networks or sequence-based autoencoders may further improve detection of advanced threats that unfold gradually over time. Furthermore, enhancing model interpretability through explainable artificial intelligence techniques could increase transparency and trust, enabling security analysts to better understand the rationale behind anomaly alerts. Finally, integration of the proposed framework with broader cloud security architectures, including security information and event management systems and automated response mechanisms, could significantly enhance its practical impact.In conclusion, this research confirms that machine learning–based anomaly detection constitutes a robust, scalable, and adaptive solution for enhancing cloud network security. By combining unsupervised learning, comprehensive evaluation, and ethical design considerations, the proposed framework addresses key limitations of traditional security approaches and provides a strong foundation for future advancements in intelligent cloud protection systems. The findings contribute meaningfully to the field of cloud security and support the broader adoption of data-driven, adaptive, and resilient security mechanisms in modern cloud computing environments.

REFERENCES

[1] Al-Mazrawe, A., & Al-Musawi, B. (2024). Anomaly detection in cloud networks: A comprehensive review of machine learning techniques. BIO Web of Conferences, 61, 02018. https://doi.org/10.1051/bioconf/20246102018
[2] Jahani, A. (2025). Machine learning-based anomaly detection in cloud computing workloads. Journal of Network and Systems Management, 33(1), 1–27. https://doi.org/10.1007/s10922-025-09701-3
[3] Baimukhanov, S., & Yessenov, K. (2025). Enhancing anomaly detection in large-scale cloud data ecosystems using deep learning. Expert Systems with Applications, 236, 121123. https://doi.org/10.1016/j.eswa.2024.121123
[4] Almajed, H., Alsaqer, A., & Albuali, A. (2025). Towards effective anomaly detection in cloud computing environments using machine learning. International Journal of Advanced Computer Science and Applications, 16(1), 442–451.
[5] Vibhute, A. D., & Patil, M. S. (2024). Deep learning-based network anomaly detection and classification: A cloud security perspective. Procedia Computer Science, 227, 341–350. https://doi.org/10.1016/j.procs.2024.01.041
[6] Kumar, D., & Yadav, J. S. (2024). A systematic review of machine learning approaches for anomaly detection in cloud computing. International Journal of Scientific Research in Science and Technology, 11(2), 89–98.
[7] Nwachukwu, C., Durodola-Tunde, K., & Akwiwu-Uzoma, C. (2024). AI-driven anomaly detection in cloud computing environments. International Journal of Science and Research Archive, 11(1), 215–225.
[8] Schummer, P. (2024). Machine learning-based network anomaly detection for cloud infrastructures. Systems, 12(3), 96. https://doi.org/10.3390/systems12030096
[9] Demirbaga, U. (2025). Advancing anomaly detection in cloud environments with generative deep learning models. Expert Systems, 42(2), e13215. https://doi.org/10.1111/exsy.13215
[10] Ahmed, Q. O., & Hasan, R. (2024). Comparative analysis of machine learning techniques for intrusion detection in cloud environments. Journal of Cloud Computing, 13(1), 77. https://doi.org/10.1186/s13677-024-00458-2
[11] Zhang, Y., Chen, H., & Li, X. (2025). Deep autoencoder-based anomaly detection for cloud network traffic. IEEE Access, 13, 21456–21468. https://doi.org/10.1109/ACCESS.2025.3342197
[12] Anomaly detection in large-scale cloud systems using unsupervised learning. (2024). arXiv preprint arXiv:2403.01892. https://arxiv.org/abs/2403.01892
[13] Artificial intelligence-based multiscale temporal modeling for anomaly detection in cloud services. (2025). arXiv preprint arXiv:2501.06741. https://arxiv.org/abs/2501.06741
[14] Contrastive learning-based dependency modeling for cloud service anomaly detection. (2025). arXiv preprint arXiv:2502.03188. https://arxiv.org/abs/2502.03188
[15] Dilworth, R., & Gudla, C. (2024). Harnessing PU learning for enhanced cloud-based DDoS anomaly detection. arXiv preprint arXiv:2405.07422. https://arxiv.org/abs/2405.07422
[16] Vacca, J. R. (2025). Cloud computing security: Foundations and challenges (2nd ed.). CRC Press.
[17] Wang, J., Liu, S., & Zhang, Q. (2025). Autoencoder-driven anomaly detection for secure cloud networking. Future Generation Computer Systems, 149, 193–205. https://doi.org/10.1016/j.future.2024.12.018
[18] AI and machine learning for cloud security: A comprehensive survey of intrusion detection systems. (2025). ACM Computing Surveys, 57(1), 1–39. https://doi.org/10.1145/3684127
[19] Cloud network anomaly detection using federated learning and explainable AI. (2025). International Journal of Security and Networks, 20(2), 145–158.
[20] Dynamic graph neural networks for early detection of cloud service anomalies. (2024). IEEE Transactions on Network and Service Management, 21(4), 4021–4034. https://doi.org/10.1109/TNSM.2024.3365981
[21] Anomaly detection using unsupervised machine learning techniques: Applications to cloud systems. (2024). Knowledge-Based Systems, 292, 111597. https://doi.org/10.1016/j.knosys.2024.111597
[22] Liu, Z., & Li, X. (2024). Error analysis of anomaly detection models under stochastic cloud traffic. Applied Mathematics and Computation, 451, 127845. https://doi.org/10.1016/j.amc.2023.127845
[23] Singh, R., & Kaur, P. (2025). Lightweight deep learning models for real-time cloud anomaly detection. Journal of Information Security and Applications, 76, 103724. https://doi.org/10.1016/j.jisa.2024.103724
[24] Chen, Y., Zhou, L., & Wang, H. (2024). Privacy-preserving anomaly detection for cloud networks using distributed learning. Computer Networks, 240, 110206. https://doi.org/10.1016/j.comnet.2024.110206
[25] Hassan, M., Rehman, S. U., & Baig, Z. A. (2025). Explainable AI for anomaly detection in cloud-based intrusion detection systems. IEEE Transactions on Dependable and Secure Computing. https://doi.org/10.1109/TDSC.2025.3354189



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY