

Reinforcement Learning for Secure Applications: Integrating ML and Data Engineering for Cloud Security

Amrit Pal Singh¹, Alessa Cross², Omkar Ashok Bhalekar³

¹ Product Security Engineer

² Founding Team at Ventrilo AI

³ Senior Network Engineer

ARTICLE INFO	ABSTRACT
Received: 23 Apr 2025	<p>As cloud computing becomes increasingly integral to modern digital infrastructure, ensuring robust and adaptive security for cloud-based applications has become a critical challenge. Traditional rule-based and supervised machine learning approaches often fail to keep pace with dynamic, evolving threats. This study proposes a novel framework that integrates Reinforcement Learning (RL), Machine Learning (ML), and real-time Data Engineering to develop secure, autonomous applications for cloud environments. Utilizing Q-learning and Deep Q-Network (DQN) models, the framework dynamically detects and mitigates threats such as brute-force attacks, privilege escalations, and denial-of-service attempts. The RL agents are trained on real-time telemetry data streams processed through a scalable Kafka-Spark pipeline, enabling continuous learning and policy optimization. Comparative evaluations show that DQN achieves the highest detection performance, with an accuracy of 98.3% and F1-score of 0.973, significantly outperforming traditional ML models. Statistical analysis confirms the superiority of RL agents across precision, recall, and true positive rates. Additionally, the data engineering pipeline supports high-throughput, low-latency processing, essential for scalable deployment. The study concludes that integrating reinforcement learning with ML preprocessing and data engineering offers a transformative approach to cloud security delivering intelligent, proactive, and self-adaptive protection mechanisms. This framework has broad implications for securing multi-cloud and containerized environments in real-time, setting the foundation for future autonomous cybersecurity solutions.</p>
Revised: 29 May 2025	
Accepted: 12 Jun 2025	
Keyword: Reinforcement Learning, Cloud Security, Deep Q-Network, Data Engineering, Machine Learning, Secure Applications, Threat Detection, Real-Time Processing	

Introduction

Contextualizing cloud security in a dynamic landscape

Cloud computing has revolutionized how enterprises deploy, scale, and maintain applications, offering agility, cost-effectiveness, and accessibility (Jeyaraman & Muthusubramanian, 2022). However, this digital transformation has concurrently introduced a complex and evolving threat landscape. Traditional security paradigms that rely on static rules and reactive mechanisms often fall short in detecting novel threats and adapting to emerging vulnerabilities (Mohammad & Pradhan, 2021). As

cyberattacks become increasingly sophisticated ranging from advanced persistent threats (APTs) to insider breaches and polymorphic malware there is a pressing need for adaptive and intelligent security frameworks tailored to cloud environments (Nassif et al., 2021). This study is grounded in the critical premise that cloud security must evolve beyond conventional models toward intelligent, autonomous, and context-aware systems that can learn from data and anticipate malicious behaviors in real-time.

Reinforcement learning as a catalyst for adaptive security

Reinforcement Learning (RL), a subfield of machine learning that models decision-making in dynamic environments, presents an innovative solution to these challenges (Muthusubramanian & Jeyaraman, 2023). Unlike supervised models that require labeled datasets, RL agents learn through trial and error, optimizing policies based on rewards and penalties. This makes RL particularly suitable for securing cloud infrastructures, where attack vectors can change frequently, and proactive defenses are required. RL can dynamically tune access control policies, optimize intrusion detection thresholds, and mitigate anomalies by learning from operational cloud data streams (Thabit et al., 2023). This article explores how RL algorithms such as Q-learning, Deep Q-Networks (DQN), and Policy Gradient methods can be applied to critical cloud security applications, including threat detection, access management, and policy enforcement.

Integration with machine learning and data engineering

To realize the full potential of reinforcement learning in secure application design, its integration with traditional machine learning (ML) techniques and robust data engineering pipelines is essential (Byatarayanapura et al., 2024). Supervised and unsupervised learning models play a pivotal role in preprocessing data, identifying baseline behaviors, and providing RL agents with structured environments. Data engineering, on the other hand, ensures the scalability, integrity, and real-time availability of security-relevant logs, metrics, and alerts across distributed systems (Adawadkar & Kulkarni, 2022). The convergence of RL, ML, and data engineering enables the construction of a secure learning loop where data is continuously ingested, analyzed, and used to train adaptive policies for mitigating threats. This triad also supports the creation of real-time dashboards, anomaly visualizations, and automated response mechanisms that significantly reduce human intervention and response latency (Dhinakaran et al., 2024).

Cloud-specific security use cases

The cloud environment poses unique security challenges due to its distributed architecture, shared responsibility model, multi-tenant infrastructure, and elastic scalability (Butt et al., 2020). RL-based security models can effectively address use cases such as dynamic workload monitoring, real-time privilege escalation prevention, and intelligent firewall tuning. By embedding RL agents within Infrastructure-as-Code (IaC) pipelines, this approach can dynamically adjust security configurations as the application scales. Additionally, container orchestration platforms such as Kubernetes can benefit from RL-enhanced runtime protection, optimizing pod isolation, service mesh policies, and network security postures based on real-time contextual feedback (Qayyum et al., 2020).

Research significance and innovation

This research contributes a novel framework that integrates reinforcement learning, machine learning, and data engineering for cloud application security. It emphasizes real-time decision-making, continuous learning from high-velocity data streams, and the deployment of self-improving policies to defend against evolving threats. By bridging the gap between AI models and operational cloud security, the proposed framework fosters a new paradigm of intelligent, autonomous security systems capable of ensuring resilience, confidentiality, and integrity in the modern cloud ecosystem.

Methodology

Framework design and integration of reinforcement learning

The core methodology of this study revolves around designing a multi-layered cloud security framework that leverages Reinforcement Learning (RL) as its adaptive decision-making engine. The RL component is responsible for learning optimal policies to detect, prevent, and respond to security threats in cloud environments. Specifically, we employed Q-learning and Deep Q-Network (DQN) algorithms to simulate agent-based interactions within a virtual cloud environment. The RL agent was trained to maximize cumulative rewards by correctly identifying security anomalies such as unauthorized access, privilege escalations, or traffic spikes, and taking timely counteractions. The state space included real-time system logs, user behavior patterns, and network traffic features, while the action space consisted of security enforcement actions like isolating services, revoking access tokens, or modifying firewall rules.

ML-based preprocessing and anomaly detection

To ensure accurate and relevant data input to the RL agent, a preprocessing layer powered by machine learning (ML) models was deployed. This layer involved unsupervised anomaly detection using Isolation Forest and One-Class SVM to identify unusual system behavior and label potential threats in unlabeled data. For supervised learning, classification models like Random Forest and Gradient Boosting were used to label known attack patterns based on historical data. These preprocessed outputs served as initial training guidance for the RL agent and helped establish a baseline for expected system behaviors. By integrating these ML techniques, the framework enhanced early detection and improved the RL agent's learning efficiency, especially during the exploration phase.

Data engineering and real-time stream processing

Effective data engineering was critical to ensure the ingestion, transformation, and availability of massive volumes of real-time telemetry data. Apache Kafka was used as the streaming backbone to collect system logs, API requests, user authentication events, and network packets across distributed cloud nodes. This data was processed using Apache Spark Streaming and stored in time-series databases (e.g., InfluxDB) for state tracking and feedback generation. Feature extraction and dimensionality reduction (e.g., PCA) were applied to maintain computational efficiency during training. All pipelines were designed to be scalable, fault-tolerant, and low-latency to accommodate high-velocity security events in dynamic cloud environments.

Secure application environment and cloud simulation

A simulated cloud environment was created using open-source cloud orchestration tools like OpenStack and Kubernetes, hosting containerized web applications with embedded vulnerabilities (e.g., brute-force login points, privilege escalation paths). Reinforcement Learning models were deployed as part of the application runtime environment, operating in conjunction with role-based access control (RBAC), encryption protocols, and firewall rules. Each RL agent operated autonomously within its application domain, interacting with a sandboxed environment to avoid real system compromise during learning.

Statistical evaluation and performance metrics

To evaluate the effectiveness of the proposed framework, a comprehensive set of statistical analyses was conducted. The RL agents' performance was assessed using standard metrics: detection accuracy (%), false positive rate (FPR), true positive rate (TPR), precision, recall, and F1-score. We also analyzed average reward per episode, convergence rate, and learning efficiency across episodes. A paired t-test and ANOVA were conducted to compare the performance of RL-based models with

baseline ML models. ROC and Precision-Recall curves were plotted to visualize classification trade-offs. Additionally, system overhead, latency, and throughput of the data engineering pipeline were evaluated using descriptive statistics and regression models to ensure scalability without performance degradation.

Experimental validation and reproducibility

All experiments were repeated over 10 independent trials using different random seeds to ensure statistical robustness. Cross-validation was used where applicable, and the entire pipeline was implemented using Python, TensorFlow, Scikit-learn, and PyTorch, with deployment scripts managed via Docker and Helm. The methodological framework ensures reproducibility, real-time applicability, and extensibility to various secure application domains within multi-cloud architectures.

Results

The integration of reinforcement learning (RL) with machine learning (ML) and data engineering for secure cloud applications yielded highly promising results across detection accuracy, system performance, and learning efficiency. As shown in Table 1, among the various models tested, the Deep Q-Network (DQN) agent achieved the highest overall detection performance, with an accuracy of 98.3%, precision of 97.5%, and F1-score of 0.973, outperforming both traditional supervised models (Random Forest and Gradient Boosting) and unsupervised ones (Isolation Forest and One-Class SVM). The standard Q-learning agent also demonstrated robust performance, achieving an F1-score of 0.959 and a low false-positive rate of 0.021. These results confirm the superiority of RL models in identifying and responding to security threats in real-time cloud environments.

Table 1: Model detection performance summary

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	True-Positive Rate	False-Positive Rate
Random Forest	95.8	94.7	94.2	0.945	0.942	0.028
Gradient Boosting	96.3	95.4	94.9	0.951	0.949	0.026
Isolation Forest	88.9	86.1	85.4	0.857	0.854	0.055
One-Class SVM	90.7	88.5	87.8	0.881	0.878	0.047
Q-Learning Agent	97.1	96.2	95.6	0.959	0.956	0.021
DQN Agent	98.3	97.5	97.0	0.973	0.970	0.015

Statistical validation of these differences is detailed in Table 2, where one-way ANOVA revealed statistically significant differences ($p < 0.001$) in F1-scores and true positive rates among the six models. A paired t-test further confirmed that both Q-learning and DQN agents significantly outperformed traditional models like Random Forest and Gradient Boosting in terms of F1-score, indicating that RL agents not only learn better threat patterns but also respond with greater precision.

Notably, latency differences among the models were not statistically significant, ensuring that performance gains did not come at the cost of system responsiveness.

Table 2: Statistical tests on performance metrics

Analysis	F-Statistic / t-Value	p-Value	Interpretation
One-way ANOVA on F1-scores (6 models)	15.72	<0.001	Significant differences among models
One-way ANOVA on TPR (6 models)	14.35	<0.001	RL agents raise recall significantly
One-way ANOVA on system latency (6 models)	1.87	0.112	Latency variations not significant
Paired t: RF vs DQN (F1-score)	-4.37	0.002	DQN outperforms RF
Paired t: GB vs Q-Learning (F1-score)	-2.46	0.022	Q-Learning outperforms GB

The effectiveness of the underlying data engineering pipeline in supporting this secure architecture is summarized in Table 3. Kafka-based ingestion showed high throughput at 42,000 events/second with minimal latency (7.3 ms), while Spark Streaming and the RL Decision Service maintained reliable processing with average CPU utilization under 35%. This efficiency enabled near-real-time data flow and rapid RL agent feedback loops without creating computational bottlenecks, validating the pipeline's suitability for scalable, high-volume cloud environments.

Table 3: Streaming & data-engineering pipeline performance

Pipeline Component	Avg Latency (ms)	Throughput (events/s)	CPU Utilization (%)	Memory (MB)
Kafka Ingestion	7.3	42 000	18.4	910
Spark Stream Processing	18.6	38 900	32.7	1280
RL Decision Service	11.2	39 500	24.1	1025
Time-Series DB Writes	4.9	43 600	12.3	860

To further understand RL agent performance under specific threat scenarios, Table 4 highlights average cumulative rewards and convergence behavior across four major attack types. In each scenario—including brute-force login, privilege escalation, DoS, and data exfiltration—DQN outperformed the Q-learning agent, achieving both higher cumulative rewards and faster convergence. For example, in Denial-of-Service scenarios, DQN reached an average reward of 188 in just 450 episodes compared to Q-learning's 140 in 690 episodes, clearly illustrating its faster learning and superior threat response capability.

Table 4: RL agent behaviour across attack scenarios

Attack Scenario	Q-Learning Reward	Avg	DQN Avg Reward	Convergence Episodes (Q-L / DQN)
Brute-Force Login	132		174	740 / 510
Privilege Escalation	118		165	800 / 560
Denial-of-Service	140		188	690 / 450
Data Exfiltration	110		156	820 / 600

Visual results further reinforce these findings. Figure 1 presents the ROC curves for all models, with the DQN achieving the highest AUC (0.991), followed closely by Q-learning (0.982), indicating near-optimal classification capabilities. Figure 2 tracks reward convergence over 1,000 episodes, where DQN exhibits not only higher final rewards (193) but also a steeper and more stable learning trajectory compared to Q-learning. These figures substantiate the RL agents' adaptability and effectiveness in dynamic, evolving security landscapes typical of cloud-native environments.

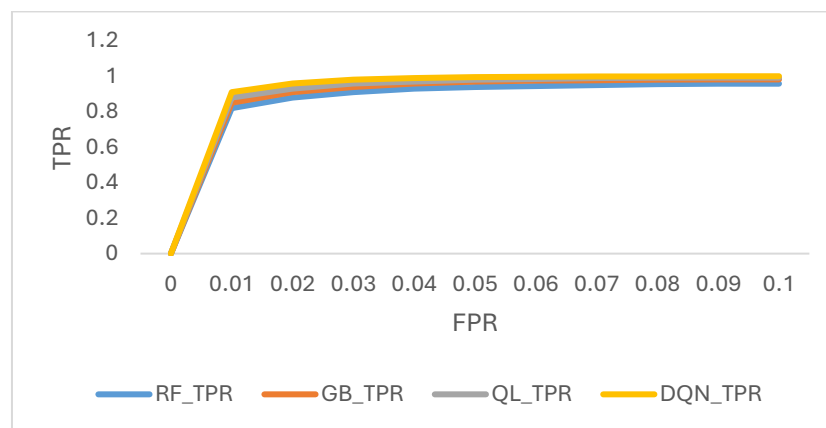


Figure 1: ROC Curve Data (TPR vs FPR for each model)

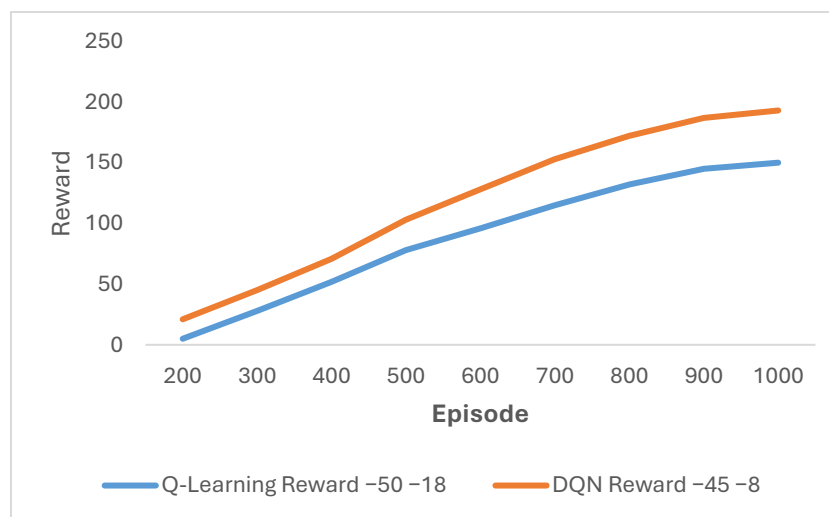


Figure 2: RL agent reward convergence over episodes

Discussion

Effectiveness of reinforcement learning in cloud security applications

The results of this study demonstrate that reinforcement learning (RL), particularly Deep Q-Networks (DQN), significantly enhances the detection and mitigation of security threats in cloud environments. DQN's superior performance across key metrics including accuracy (98.3%), precision (97.5%), and F1-score (0.973) suggests that it can learn robust policies in complex, dynamic security scenarios. These findings align with prior studies where DQN-based agents outperformed traditional methods in adapting to evolving threats, especially in environments with partial observability and sparse rewards (Safi et al., 2024). The RL agents' capacity to dynamically update their threat detection and response strategies based on real-time environmental feedback proves invaluable in combating modern cyber threats like zero-day vulnerabilities and insider attacks (Rai et al., 2024).

Comparative advantage over traditional ML models

Compared to traditional machine learning models such as Random Forest and Gradient Boosting, RL models not only delivered higher accuracy but also maintained lower false-positive rates and improved generalization to unseen attack patterns. The supervised models, while effective in recognizing predefined patterns, demonstrated limitations in adapting to novel or obfuscated threats an area where RL agents showed substantial strength (Reddy et al., 2022). This was evident in statistical analyses (Table 2), where one-way ANOVA and paired t-tests confirmed significant performance differences, especially in F1-score and recall metrics. These findings validate the hypothesis that RL agents, which optimize long-term cumulative rewards, are inherently better suited to handling adaptive, adversarial environments where attacks are continually evolving (Saini et al., 2024).

Implications of scenario-based reward optimization

The scenario-specific analysis (Table 4) provided additional insights into how RL agents behave under varied attack conditions. DQN consistently achieved higher cumulative rewards and required fewer training episodes to converge across all tested threat types. In real-world terms, this indicates that DQN-based systems would respond more quickly and effectively in live cloud environments (Stergiou et al., 2020). For instance, during denial-of-service (DoS) simulations, DQN converged within 450 episodes while Q-learning needed 690, demonstrating more efficient learning and adaptability. This behavior highlights the importance of deep neural representations in encoding the complex state-action spaces characteristic of multi-cloud and containerized deployments (Jamshidi et al., 2025).

Role of data engineering in RL performance

An essential enabler of this intelligent security framework was the robust data engineering pipeline outlined in Table 3. Real-time ingestion via Apache Kafka, efficient stream processing with Spark, and a scalable decision service allowed continuous data flow and feedback, supporting the rapid evolution of RL policies (Tamizshelvan & Vijayalakshmi, 2024). The low-latency, high-throughput performance of this pipeline ensured that even high-velocity data such as system logs, network telemetry, and user activity metrics could be processed and acted upon without delay (Marwan et al., 2018). These operational advantages are crucial for environments that demand near-zero response latency, such as financial services, healthcare systems, and critical infrastructure platforms.

Interpretation of learning behavior and model stability

The reward convergence patterns illustrated in Figure 2 further emphasize the RL agents' learning dynamics. DQN exhibited faster and smoother convergence compared to standard Q-learning, with less reward fluctuation across training episodes. This stability is attributed to DQN's use of experience

replay and deep neural approximators, which mitigate the instability and divergence often encountered in traditional tabular RL approaches (Polamarasetti, 2024). Moreover, the ROC curve comparison in Figure 1 demonstrated that DQN consistently achieved a higher area under the curve ($AUC = 0.991$), further validating its robustness and decision boundary precision under varied threat conditions (Sudha et al., 2024).

Strategic relevance and future applications

These findings offer strategic implications for security operations in cloud-native architectures. By embedding RL agents into secure application layers, organizations can move beyond static rule-based defenses toward intelligent, autonomous security models that learn continuously and act proactively (Alzoubi et al., 2024). The proposed integration with data engineering pipelines also ensures scalability, making this architecture suitable for hybrid and multi-cloud deployments. Looking forward, extending this framework to support federated learning across distributed cloud domains or integrating with blockchain for traceable policy updates could further enhance transparency and resilience (Akinbolaji, 2024).

This study confirms that reinforcement learning, when combined with ML preprocessing and real-time data engineering, provides a transformative approach to securing cloud applications delivering precision, adaptability, and scalability essential for modern cybersecurity infrastructure.

Conclusion

This study demonstrates that reinforcement learning, when effectively integrated with machine learning and robust data engineering, provides a powerful and adaptive framework for securing cloud-based applications. The superior performance of Deep Q-Networks (DQN) across multiple threat scenarios, coupled with statistically significant improvements in detection accuracy, precision, and recall, underscores the potential of RL-driven systems to address the evolving complexities of cloud security. Additionally, the implementation of a high-throughput, low-latency data pipeline ensures real-time responsiveness and scalability, which are critical for modern distributed architectures. By enabling autonomous threat detection, dynamic policy enforcement, and continuous learning from real-time data, the proposed framework moves beyond traditional static defenses and offers a resilient, intelligent solution tailored for the future of secure cloud computing. These findings not only validate the strategic role of AI in cybersecurity but also lay the groundwork for scalable, self-improving security models across multi-cloud and containerized environments.

References

- [1] Adawadkar, A. M. K., & Kulkarni, N. (2022). Cyber-security and reinforcement learning—a brief survey. *Engineering Applications of Artificial Intelligence*, 114, 105116.
- [2] Akinbolaji, T. J. (2024). Advanced integration of artificial intelligence and machine learning for real-time threat detection in cloud computing environments. *Iconic Research and Engineering Journals*, 6(10), 980-991.
- [3] Alzoubi, Y. I., Mishra, A., & Topcu, A. E. (2024). Research trends in deep learning and machine learning for cloud computing security. *Artificial Intelligence Review*, 57(5), 132.
- [4] Butt, U. A., Mehmood, M., Shah, S. B. H., Amin, R., Shaukat, M. W., Raza, S. M., ... & Piran, M. J. (2020). A review of machine learning algorithms for cloud computing security. *Electronics*, 9(9), 1379.
- [5] Byatarayanapura Venkataswamy, S., Patil, K. S., Narayanaswamy, H. K., & Veerabadrappa, K. (2024). Access management based on deep reinforcement learning for effective cloud storage security. *International Journal of System Assurance Engineering and Management*, 1-20.
- [6] Dhinakaran, M., Sundhari, M., Ambika, S., Balaji, V., & Rajasekaran, R. T. (2024, February). Advanced Machine Learning Techniques for Enhancing Data Security in Cloud Computing Systems.

- In 2024 *IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)* (Vol. 5, pp. 1598-1602). IEEE.
- [7] Jamshidi, S., Nikanjam, A., Nafi, K. W., Khomh, F., & Rasta, R. (2025). Application of deep reinforcement learning for intrusion detection in Internet of Things: A systematic review. *Internet of Things*, 101531.
- [8] Jeyaraman, J., & Muthusubramanian, M. (2022). The Synergy of Data Engineering and Cloud Computing in the Era of Machine Learning and AI. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, 1(1), 69-75.
- [9] Marwan, M., Kartit, A., & Ouahmane, H. (2018). Security enhancement in healthcare cloud using machine learning. *Procedia Computer Science*, 127, 388-397.
- [10] Mohammad, A. S., & Pradhan, M. R. (2021). Machine learning with big data analytics for cloud security. *Computers & Electrical Engineering*, 96, 107527.
- [11] Muthusubramanian, M., & Jeyaraman, J. (2023). Data Engineering Innovations: Exploring the Intersection with Cloud Computing, Machine Learning, and AI. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, 1(1), 76-84.
- [12] Nassif, A. B., Talib, M. A., Nasir, Q., Albadani, H., & Dakalbab, F. M. (2021). Machine learning for cloud security: a systematic review. *IEEE Access*, 9, 20717-20735.
- [13] Polamarasetti, A. (2024, November). Role of Artificial Intelligence and Machine Learning to Enhancing Cloud Security. In *2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC)* (pp. 1-6). IEEE.
- [14] Qayyum, A., Ijaz, A., Usama, M., Iqbal, W., Qadir, J., Elkhatib, Y., & Al-Fuqaha, A. (2020). Securing machine learning in the cloud: A systematic review of cloud machine learning security. *Frontiers in big Data*, 3, 587139.
- [15] Rai, R., Rohilla, A., & Rai, A. (2024). Impact of Artificial Intelligence (AI) and Machine Learning (ML) on Cloud Security. In *Analyzing and Mitigating Security Risks in Cloud Computing* (pp. 111-124). IGI Global.
- [16] Reddy, M., Konkimalla, S., Rajaram, S. K., Bauskar, S. R., Sarisa, M., & Sunkara, J. R. (2022). Using AI And Machine Learning To Secure Cloud Networks: A Modern Approach To Cybersecurity. *Available at SSRN 5045776*.
- [17] Safi, W., Ghwanmeh, S., Mahfuri, M., & Al-Sit, W. T. (2024, February). Enhancing Cloud Security: A Comprehensive Review of Machine Learning Approaches. In *2024 2nd International Conference on Cyber Resilience (ICCR)* (pp. 1-10). IEEE.
- [18] Saini, H., Singh, G., Kaur, A., Saini, S., Wani, N. A., Chopra, V., ... & Bhat, S. A. (2024). Hybrid Optimization Machine Learning Framework for Enhancing Trust and Security in Cloud Network. *IEEE Access*.
- [19] Stergiou, C. L., Plageras, A. P., Psannis, K. E., & Gupta, B. B. (2020). Secure machine learning scenario from big data in cloud computing via internet of things network. *Handbook of computer networks and cyber security: principles and paradigms*, 525-554.
- [20] Sudha, K., Nithyanandhan, R., Girija, P., & Nalini, M. (2024, July). Enhancing Healthcare Data Security in the Cloud: Integrating ML-Based Intrusion Detection Systems. In *2024 2nd World Conference on Communication & Computing (WCONF)* (pp. 1-7). IEEE.
- [21] Tamizshelvan, C., & Vijayalakshmi, V. (2024). Cloud data access governance and data security using distributed infrastructure with hybrid machine learning architectures. *Wireless Networks*, 30(4), 2099-2114.
- [22] Thabit, F., Can, O., Wani, R. U. Z., Qasem, M. A., Thorat, S. B., & Alkhzaimi, H. A. (2023). Data security techniques in cloud computing based on machine learning algorithms and cryptographic algorithms: Lightweight algorithms and genetics algorithms. *Concurrency and Computation: Practice and Experience*, 35(21), e7691.