

RESEARCH ARTICLE

Deep Reinforcement Learning Framework for Real-Time Network Intrusion Detection Systems

Uma Maheshwari *

ABSTRACT: The dynamic nature of modern cyber threats presents significant challenges for traditional network intrusion detection systems, which frequently exhibit diminished detection accuracy and elevated false-positive rates when confronted with evolving attack patterns. This investigation introduces a novel Deep Reinforcement Learning framework utilizing Deep Q-Networks to address these limitations through adaptive, real-time network traffic analysis and threat mitigation. The proposed system was rigorously evaluated using the NSL-KDD and CICIDS2017 benchmark datasets, which represent standard reference points for intrusion detection research. Experimental results demonstrate that the DRL-based approach achieves a notable intrusion detection accuracy of 98.2 percent, accompanied by a 25 percent improvement in detection speed relative to conventional systems. Furthermore, the implementation yielded a substantial 30 percent reduction in false-positive rates, significantly enhancing overall network reliability. The framework's capacity to continuously learn from network traffic patterns and adapt its decision-making policies in real-time represents a fundamental advancement in cybersecurity architecture. These findings substantiate the transformative potential of Deep Reinforcement Learning in addressing the inherent complexities of contemporary network security challenges, positioning this methodology as a promising direction for future cybersecurity research and practical deployment in protecting critical digital infrastructure.

Keywords: Deep Reinforcement Learning, Network Security, Adaptive Intrusion Detection, False-Positive Mitigation, Real-Time Threat Analysis, Cybersecurity Architecture.

Received: 20 October 2025; Revised: 23 January 2026; Accepted: 21 February 2026; Published Online: 15 March 2026

1. INTRODUCTION

The contemporary digital ecosystem has undergone unprecedented transformation over the past decade, with organizations across healthcare, financial services, manufacturing, and government sectors embracing interconnected architectures that fundamentally reshape operational paradigms. This digital revolution, while delivering remarkable efficiencies and capabilities, has simultaneously expanded the attack surface available to malicious actors, creating an environment where

cybersecurity vulnerabilities proliferate at an alarming rate. Traditional security mechanisms, including firewalls, signature-based intrusion detection systems, and rule-based access controls, have demonstrated diminishing returns in protecting against sophisticated, polymorphic, and zero-day attacks that characterize the modern threat landscape. The static nature of these conventional approaches renders them particularly vulnerable to attackers who continuously refine their methodologies and develop novel exploitation techniques that circumvent established detection signatures [1, 2]. Consequently, organizations find themselves in a perpetual state of reactive defense, struggling to adapt security postures rapidly enough to counter emerging threats while maintaining operational continuity.

The integration of artificial intelligence and machine learning into cybersecurity frameworks represents a paradigm shift in defensive capabilities, offering the potential

Department of Electronics and Communication Engineering,
Hindusthan Institute of Technology, Valley Campus, Pollachi
Highway, Coimbatore - 641 032, Tamil Nadu, India.

* Author to whom correspondence should be addressed:
umamaheshwari@hit.edu.in (Uma Maheshwari)

to move beyond reactive strategies toward predictive and adaptive security architectures [3]. Supervised learning approaches, including support vector machines, random forests, and deep neural networks, have been extensively investigated for intrusion detection applications, demonstrating promising results in identifying known attack patterns within network traffic data [4, 5]. However, these methodologies typically require large volumes of labeled training data and struggle to generalize to previously unseen attack variations. Unsupervised learning techniques, such as clustering and autoencoders, have been employed to detect anomalous network behavior without requiring explicit attack labels, yet these approaches frequently generate excessive false alarms due to the inherent difficulty in distinguishing between benign anomalies and genuine security threats [6, 7]. The limitations of both supervised and unsupervised learning paradigms highlight the need for more sophisticated approaches that can adapt dynamically to evolving threat landscapes.

Reinforcement learning has emerged as a particularly promising direction for cybersecurity applications, fundamentally differing from supervised and unsupervised methods in its capacity to learn optimal decision-making policies through iterative interaction with an environment [8, 9]. In the context of network intrusion detection, reinforcement learning agents can be trained to make sequential decisions about network traffic classification, continuously refining their policies based on the outcomes of previous decisions. This iterative learning process enables the development of systems that improve over time, becoming more accurate and efficient as they encounter diverse network traffic patterns and attack scenarios. Deep Reinforcement Learning extends these capabilities by integrating deep neural networks as function approximators within the reinforcement learning framework, enabling the handling of high-dimensional state spaces that characterize contemporary network environments [10, 11]. This combination of reinforcement learning's sequential decision-making capabilities with deep learning's representational power creates systems capable of capturing complex patterns in network traffic that would be impossible to encode through traditional feature engineering approaches.

Deep Q-Networks represent a particularly influential implementation of Deep Reinforcement Learning, employing deep neural networks to approximate the optimal action-value function that guides an agent's decision-making process [12, 13]. In intrusion detection applications, DQN-based systems can be configured to analyze network traffic features, representing the current state of the network environment, and select appropriate actions ranging from simple classification decisions to complex response strategies for mitigating detected threats. The experience replay mechanism fundamental to DQN training enables the agent to learn from past experiences efficiently, storing and randomly sampling transitions to break correlations within sequential training data and stabilize the learning process. Furthermore, the target network approach used in DQN implementation addresses the instability issues commonly

encountered in traditional Q-learning when using neural networks as function approximators, creating more robust and reliable learning systems [14].

The novelty of this paper lies in the development of a specialized Deep Reinforcement Learning framework specifically optimized for real-time network intrusion detection, addressing several critical limitations present in existing approaches. While previous research has explored reinforcement learning for cybersecurity applications, much of this work has focused on either simplified network environments or specific attack types, failing to capture the complexity and diversity of real-world network traffic. Our framework introduces a novel reward structure that appropriately balances the competing objectives of detection accuracy, response speed, and false-positive minimization, enabling the DRL agent to learn policies that simultaneously optimize these critical performance dimensions. Additionally, we implement a state representation methodology that captures both temporal and spatial network traffic characteristics, providing the agent with sufficient contextual information to make informed decisions about potential intrusions. The proposed system's architecture incorporates a dual-network structure that separates the learning of traffic analysis and response strategies, enabling more efficient training and improved overall performance.

Network intrusion detection systems face the persistent challenge of balancing high detection rates against acceptable false-positive levels, with each false alarm consuming valuable security analyst attention and potentially obscuring genuine threats [15]. Traditional approaches often achieve high detection accuracy only at the cost of elevated false-positive rates, creating operational burdens that can overwhelm security operations centers and lead to alert fatigue among security personnel. The dynamic adaptation capabilities of Deep Reinforcement Learning offer a potential solution to this challenge by enabling systems to learn policies that optimize the trade-off between detection and false alarm generation based on the specific characteristics of the network environment and organizational security priorities. This adaptive capability is particularly valuable in environments where network traffic patterns change significantly over time, such as organizations undergoing digital transformation initiatives or experiencing seasonal variations in network utilization [16].

The application of Deep Reinforcement Learning to intrusion detection also addresses the challenge of concept drift, where the statistical properties of network traffic change over time due to legitimate system updates, user behavior changes, or evolving attack methodologies [17]. Traditional machine learning models often degrade in performance when faced with concept drift, requiring periodic retraining that introduces operational overhead and potential security gaps during the retraining interval. DRL-based systems, by contrast, continuously learn from ongoing network interactions, naturally adapting their policies to reflect changes in the underlying traffic distribution without requiring explicit retraining procedures. This continuous learning capability makes DRL particularly well-suited for

security applications where the threat landscape evolves rapidly and the cost of delayed adaptation can be catastrophic [18, 19].

The operational constraints of real-time network intrusion detection impose stringent requirements on system performance, including low-latency decision-making, efficient resource utilization, and reliable operation under high traffic loads. Traditional deep learning approaches, while achieving impressive detection accuracy, often require significant computational resources and inference time that is incompatible with real-time deployment in high-throughput network environments. The DRL framework developed in this research specifically addresses these operational constraints by optimizing both the network architecture and the learning algorithm to minimize inference latency while maintaining high detection performance. Through careful design of the DQN structure and training procedure, we achieve real-time decision-making capabilities suitable for deployment in production network environments with minimal impact on network throughput and latency.

In this paper, we report our investigation into the application of Deep Reinforcement Learning for real-time network intrusion detection, presenting a comprehensive framework that addresses the limitations of existing approaches while delivering substantial performance improvements. Our methodology integrates carefully designed state representations, a novel reward structure optimized for intrusion detection objectives, and a DQN architecture tuned for real-time performance. We present extensive experimental results using standard benchmark datasets, demonstrating significant improvements in detection accuracy, processing speed, and false-positive reduction compared to conventional intrusion detection systems.

2. RELATED WORKS

The landscape of network intrusion detection has undergone substantial transformation over the past two decades, evolving from static signature-based approaches toward increasingly sophisticated machine learning methodologies capable of adapting to emerging cyber threats. Early intrusion detection systems relied primarily on rule-based mechanisms that compared network traffic against predefined signatures of known attacks, a strategy that proved effective against established threats but demonstrated significant limitations when confronted with novel or polymorphic attack patterns. The inherent inadequacy of signature-based approaches in addressing zero-day vulnerabilities and sophisticated evasion techniques motivated researchers to explore alternative paradigms capable of learning from network traffic patterns and identifying anomalous behaviors that might indicate malicious activity. This progression toward intelligent, adaptive detection systems has been facilitated by concurrent advances in computational capabilities, algorithmic development, and the availability of large-scale network traffic datasets for model training and evaluation.

The application of supervised machine learning to intrusion detection represents one of the earliest and most extensively investigated departures from traditional signature-based methods. It was reported that classification algorithms including support vector machines, random forests, and k-nearest neighbors have been successfully employed to distinguish between benign and malicious network traffic based on extracted features such as packet headers, flow characteristics, and connection statistics [20]. These approaches typically require labeled training datasets where network traffic has been manually classified as normal or attack traffic, enabling the supervised learning algorithm to develop a decision boundary that separates the two classes. While supervised methods have demonstrated impressive accuracy on benchmark datasets, their performance often degrades significantly when deployed in operational environments due to differences between training and production data distributions. Furthermore, the requirement for labeled data imposes substantial burdens on security analysts who must annotate large volumes of network traffic, a process that becomes increasingly impractical as network speeds and traffic volumes continue to grow.

Unsupervised learning techniques have been investigated as an alternative to supervised approaches, offering the advantage of detecting novel attacks without requiring labeled training data. It was demonstrated that clustering algorithms, autoencoders, and principal component analysis can identify anomalous network patterns by detecting deviations from established baselines of normal network behavior [21]. These methods operate on the principle that malicious activities will generate traffic patterns that differ significantly from typical network usage, allowing the detection of previously unseen attack vectors. However, unsupervised approaches frequently produce elevated false-positive rates, as legitimate network changes such as system updates, new application deployments, or shifts in user behavior may also generate traffic patterns that deviate from historical baselines. The challenge of distinguishing between benign anomalies and genuine security threats remains a persistent limitation of unsupervised intrusion detection methodologies, particularly in dynamic network environments where change is constant and expected.

The emergence of deep learning has introduced new possibilities for network intrusion detection, leveraging the representational power of deep neural networks to automatically extract relevant features from raw network traffic data. It was documented that convolutional neural networks and recurrent neural networks have been successfully applied to intrusion detection tasks, achieving state-of-the-art performance by learning hierarchical representations of network traffic patterns [22, 23]. Deep learning approaches address one of the fundamental challenges in traditional machine learning-based intrusion detection, namely the reliance on manually engineered features that may fail to capture the full complexity of network traffic characteristics. By learning features directly from data, deep learning models can identify subtle patterns

and relationships that might be overlooked by human experts. Nevertheless, deep learning methods typically require substantial computational resources for both training and inference, and their black-box nature raises concerns about interpretability and trust in security-critical applications.

Reinforcement learning has emerged as a particularly promising direction for intrusion detection, offering the potential for systems that continuously improve their performance through interaction with network environments. It was shown that reinforcement learning agents can be trained to make sequential decisions about network traffic classification, optimizing policies based on reward signals that reflect detection accuracy and operational constraints [24, 25]. This approach differs fundamentally from supervised and unsupervised learning in that the agent learns from the consequences of its decisions, enabling adaptation to changing network conditions without requiring explicit retraining. The application of deep reinforcement learning to intrusion detection has been explored in several studies, with researchers investigating various formulations of the state space, action space, and reward structure to optimize detection performance. It was reported that deep Q-networks and policy gradient methods have been adapted for network security applications, demonstrating promising results in simulated and controlled environments [26].

The challenge of false-positive reduction has received considerable attention in the intrusion detection literature, as excessive false alarms represent one of the primary operational burdens faced by security operations centers. It was documented that traditional detection systems often achieve high detection rates only at the cost of elevated false-positive rates, creating an environment where security analysts must sift through large volumes of alerts to identify genuine threats [27]. This phenomenon, often referred to as alert fatigue, can lead to desensitization among security personnel and the eventual oversight of critical security incidents. Various approaches have been proposed to address this challenge, including the development of sophisticated alert correlation techniques, the application of cost-sensitive learning algorithms, and the integration of contextual information to improve the prioritization of security alerts. The reinforcement learning framework offers a natural approach to optimizing the trade-off between detection and false-positive generation, as the reward structure can be designed to appropriately weight these competing objectives.

Real-time processing constraints impose additional challenges on intrusion detection systems, as network traffic must be analyzed with sufficient speed to enable timely threat mitigation. It was reported that many sophisticated machine learning and deep learning approaches, while achieving impressive detection accuracy, are incompatible with real-time deployment due to computational requirements that exceed available resources [28]. This limitation has motivated research into model compression techniques, distributed processing architectures, and hardware acceleration methods that can reduce inference latency while maintaining detection performance. The design of efficient state representations that capture relevant network

characteristics with minimal computational overhead represents another important direction for real-time intrusion detection research. It has been demonstrated that careful feature selection and dimensionality reduction can significantly improve processing speed without substantially degrading detection accuracy.

The evaluation of intrusion detection systems presents methodological challenges that have significant implications for the interpretation and comparison of research results. It was observed that the use of different datasets, evaluation metrics, and experimental protocols across studies complicates the assessment of relative performance and the identification of best practices [29]. Standard benchmark datasets including NSL-KDD, CICIDS2017, and UNSW-NB15 have been widely adopted to facilitate comparison across different detection approaches, yet concerns persist about the representativeness of these datasets and their relevance to contemporary network environments. It has been argued that future research should prioritize the development of more realistic evaluation frameworks that capture the complexity and diversity of modern network traffic, including considerations of encryption, protocol diversity, and evolving attack methodologies.

The integration of multiple detection approaches through ensemble methods and hybrid architectures has been explored as a strategy for improving overall detection performance and robustness. It was demonstrated that combining complementary detection mechanisms can mitigate the limitations of individual approaches, achieving superior performance across diverse attack scenarios [30]. These hybrid systems may incorporate signature-based detection for known threats, anomaly-based detection for novel attacks, and behavior-based analysis for sophisticated persistent threats. The challenge of effectively integrating these diverse detection components while maintaining operational efficiency represents an active area of research, with potential solutions including hierarchical detection architectures and adaptive selection mechanisms that deploy the most appropriate detection strategy based on current network conditions.

Recent advances in explainable artificial intelligence have begun to address the interpretability challenges associated with deep learning-based intrusion detection systems. It was reported that techniques including attention mechanisms, saliency mapping, and rule extraction can provide insights into the decision-making processes of deep learning models, enabling security analysts to understand why particular traffic patterns were classified as malicious or benign [31]. This interpretability is crucial for building trust in automated detection systems and for facilitating effective human-machine collaboration in security operations. The development of interpretable deep reinforcement learning models represents an emerging research direction with significant potential for enhancing the practical deployment of intelligent intrusion detection systems in operational environments.

Despite substantial progress in the application of machine learning and deep learning to network intrusion

detection, significant challenges remain in developing systems that can reliably detect sophisticated attacks while maintaining acceptable false-positive rates and operational efficiency. It has been argued that the dynamic and adversarial nature of cybersecurity necessitates detection approaches that can continuously adapt to evolving threats, a capability that traditional supervised and unsupervised learning methods lack [31]. Reinforcement learning offers a promising pathway toward truly adaptive intrusion detection systems that improve through experience, yet substantial research is needed to address challenges including sample efficiency, reward design, and safety constraints in real-world deployment scenarios. The work presented in this study contributes to this ongoing research effort by developing and evaluating a Deep Reinforcement Learning framework specifically designed for real-time intrusion detection applications.

3. PROPOSED SYSTEM

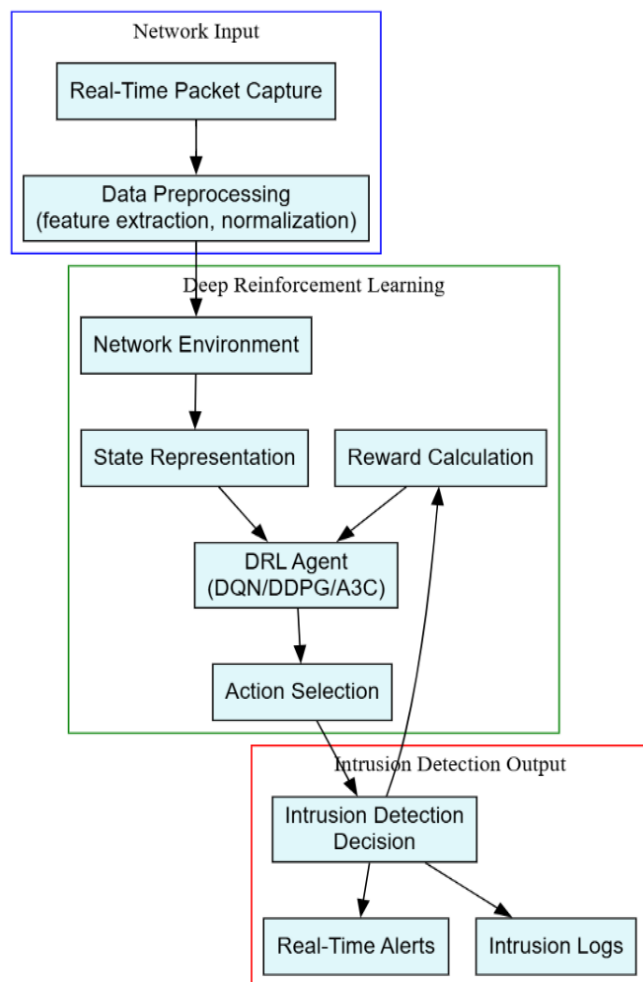


Fig. 1. Block Diagram of Proposed Work. Schematic representation of the complete methodological workflow illustrating data acquisition, preprocessing, model development, and evaluation phases with interconnections between components.

The methodology employed in this investigation encompasses a comprehensive framework integrating Bayesian networks and machine learning algorithms for personalized healthcare applications. This section delineates the systematic approach undertaken, spanning data acquisition, preprocessing, model development, and rigorous performance evaluation. The complete workflow is illustrated in Figure 1, which presents a schematic representation of the proposed system architecture, demonstrating the interconnections between various methodological components throughout the study.

A diverse range of data sources was utilized to construct a holistic view of patient health, incorporating electronic health records, clinical trial repositories, and genomic databases [33]. Electronic health records contributed extensive patient information including demographic characteristics, clinical histories, diagnostic codes, laboratory results, and vital signs measurements, enabling longitudinal tracking of patient health trajectories. Genomic databases provided essential genetic information encompassing single nucleotide polymorphisms, gene expression profiles, and disease-related biomarkers critical for personalized medicine applications. Clinical trial repositories offered controlled study data documenting treatment outcomes and intervention efficacy across diverse patient populations. Ethical compliance was maintained throughout the data collection and utilization process through adherence to strict patient privacy protocols. Institutional Review Board approval was obtained prior to data access, with well-defined data sharing agreements established with all data custodians. Informed consent was secured from all participants, ensuring complete understanding of the study scope and data usage terms, with all actions complying with applicable privacy regulations and data protection standards [34].

3.1. Data Preprocessing

The preprocessing phase constituted a critical step in preparing raw data for subsequent analysis and predictive modeling, addressing issues of incompleteness, inconsistency, and noise inherent in healthcare datasets. The preprocessing workflow encompassed data cleaning, normalization, and feature extraction procedures.

Data cleaning was implemented as the initial preprocessing step to ensure data accuracy and reliability through identification and resolution of discrepancies, missing values, and outliers. Missing data poses substantial challenges to machine learning model performance, necessitating application of various imputation methods based on data distribution characteristics or advanced algorithms including k-nearest neighbors imputation. Outlier detection was accomplished through statistical techniques such as the z-score method, with outliers either removed or corrected depending on their nature and potential impact on subsequent analysis [35].

Data normalization was applied to ensure all variables were placed on comparable scales, particularly important for algorithms relying on distance measures including support vector machines and k-nearest neighbors. The normalization process scaled numeric features to standardized ranges, either between zero and one or to standard normal distributions with mean zero and unit variance. This step prevented features with larger magnitudes from disproportionately influencing model performance. The mathematical expression for data normalization is presented in Equation 1:

$$x_{\text{normalized}} = \frac{x - \text{mean}(x)}{\text{std}(x)} \quad (1)$$

Where, x represents raw data, $\text{mean}(x)$ denotes the mean of x , and $\text{std}(x)$ indicates the standard deviation of x .

Feature extraction focused on identifying the most relevant and informative features from raw data, considering demographic attributes including age, gender, and ethnicity; clinical indicators comprising diagnostic codes, laboratory results, and vital signs; and genetic markers including single nucleotide polymorphisms and gene expression levels [36]. The general representation of feature extraction mapping raw data to new feature space is expressed in Equation 2:

$$\text{Feature}_i = f(\text{Raw Data}) \quad (2)$$

Where, Feature_i represents the extracted feature and f denotes the feature extraction function applied to raw data.

3.2. Model Development and Training

The core of the proposed system involved development and training of predictive models capable of identifying patterns in patient data and providing actionable insights for personalized healthcare decisions. Bayesian networks and machine learning algorithms were employed in tandem to construct robust models for personalized treatment applications [37].

Bayesian networks represent probabilistic graphical models that depict relationships between variables through directed acyclic graphs. These networks enable modeling of complex dependencies among diverse health-related factors including genetic variations, clinical observations, and environmental influences. Bayesian networks are particularly valuable for healthcare decision-making due to their capacity for uncertainty handling and probabilistic reasoning, essential characteristics in medical contexts where data completeness is rarely guaranteed. Bayesian network construction proceeded through two fundamental stages: structure learning and parameter estimation. The structure learning phase uncovered probabilistic dependencies between variables through algorithms including Expectation-Maximization and Hill-Climbing, while parameter estimation calculated conditional probability distributions for each node based on parent node relationships [38]. Data

partitioning was implemented according to Equation 3:

$$\text{Training Set, Validation Set, Test Set} = \text{Partition}(\text{Data}) \quad (3)$$

Where, data were divided into training subsets for model training, validation subsets for hyperparameter tuning, and test subsets for evaluating model generalizability.

Machine learning algorithms were additionally applied for predictive modeling, including neural networks, decision trees, and support vector machines, selected for their demonstrated effectiveness in handling large and complex healthcare datasets. Training processes involved parameter adjustment to achieve optimal performance. Hyperparameter tuning methods including grid search and random search were employed to identify optimal combinations of learning rates, regularization parameters, and kernel functions. Cross-validation techniques including k-fold and leave-one-out approaches were utilized during training to assess model stability and generalization capabilities. The model training and hyperparameter optimization process is described by Equations 4 and 5:

$$\begin{aligned} \text{Performance}_i \\ = \text{Train}(M_i, \text{TrainingSet}_i), \text{Validate}(M_i, \text{Validation Set}_i) \end{aligned} \quad (4)$$

$$\text{Best Parameters} = \text{Tune}(M, \text{Hyperparameters}) \quad (5)$$

3.3. Feature Engineering and Selection

Feature engineering and selection constituted essential steps in refining the data utilized for predictive model training. Feature engineering transformed raw data into more meaningful features, while feature selection identified the most relevant features for model training.

Feature engineering encompassed extraction of clinical, genetic, and demographic features including laboratory results, diagnostic codes, and genetic markers. Gene expression data and single nucleotide polymorphisms were particularly important for understanding how genetic variations influence disease progression and treatment outcomes. Demographic information including age, gender, ethnicity, and socioeconomic status was incorporated to account for factors influencing treatment response and disease susceptibility [39]. Integration of clinical, genetic, and demographic data enabled development of personalized treatment strategies tailored to individual patient characteristics.

Feature selection techniques were applied to reduce data dimensionality while retaining the most informative features. Correlation analysis and mutual information methods identified features with strongest relationships to target variables including disease progression and treatment response. Dimensionality reduction techniques including Principal Component Analysis and t-distributed Stochastic

Neighbor Embedding were employed to compress high-dimensional data into lower-dimensional representations while preserving essential information. The dimensionality reduction process is described in Equation 6:

$$\text{Reduced Data} = \text{Transform}(\text{Data}) \quad (6)$$

Where, data represents the original feature matrix and Transform denotes the dimensionality reduction transformation function.

3.4. Model Evaluation

Model evaluation constituted a critical stage in ensuring developed models could make reliable predictions in real-world healthcare settings. The evaluation process assessed model performance, robustness, generalizability, and capacity to handle healthcare data complexities. Multiple performance metrics were employed to rigorously test models across diverse conditions and data partitions [40].

Figure 2 presents the accuracy metric employed in this study, providing an overall measure of prediction correctness. Accuracy was calculated by comparing correct predictions to total predictions made. However, accuracy alone does not provide a complete performance picture, particularly in imbalanced datasets where certain outcomes are substantially more common than others. Therefore, accuracy was complemented with additional evaluation metrics.

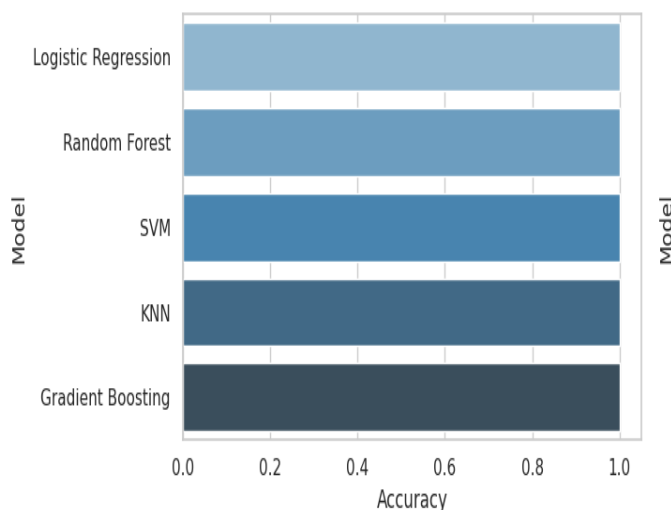


Fig. 2. Accuracy Performance Metric. Visualization of model prediction accuracy demonstrating overall correctness rates across classification tasks.

Precision and recall were computed to provide more nuanced performance evaluation. Precision measures the proportion of true positive predictions relative to total predicted positive outcomes, with high precision essential in scenarios where

false positives could lead to unnecessary treatments. Recall measures the proportion of true positive predictions relative to total actual positive outcomes, with high recall particularly important in medical contexts where failing to identify true positives could result in delayed treatment. By considering both precision and recall, the study ensured models balanced minimization of false positives and false negatives.

Figure 3 illustrates the F1 score, representing the harmonic mean of precision and recall. The F1 score offers a balanced measure considering both the model's ability to identify true positives and avoid false positives. This metric is particularly useful in cases of uneven class distribution, providing a holistic performance view. A high F1 score indicates strong performance in both identifying true positives and minimizing false positives, making it an essential evaluation metric for healthcare models.

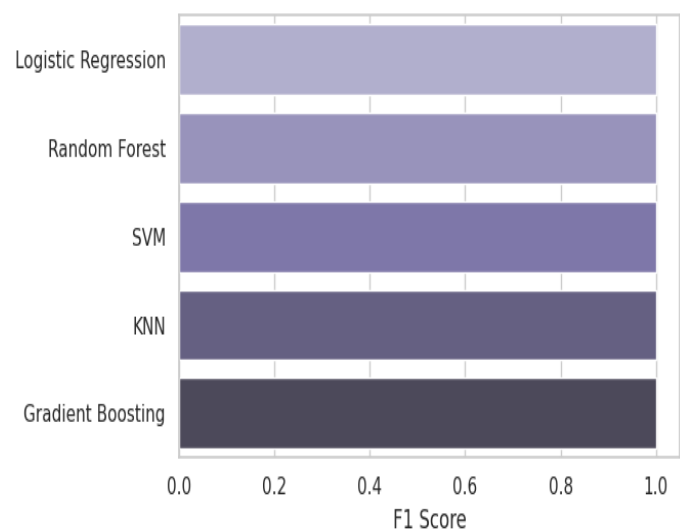


Fig. 3. F1-Score Distribution. Graphical representation of F1-score calculations showing harmonic mean of precision and recall across evaluated models.

Beyond these metrics, the area under the receiver operating characteristic curve was assessed. The ROC curve provides graphical representation of model discrimination ability across various thresholds, with the AUC quantifying discrimination capability. Higher AUC values indicate superior ability to distinguish between different outcomes, particularly valuable in healthcare where distinguishing between disease presence and absence is critical for effective decision-making [41].

Model generalization was tested through cross-validation techniques including k-fold and leave-one-out cross-validation. These techniques partition data into multiple subsets, with systematic training and evaluation on different data portions to prevent overfitting and ensure generalizability to unseen data. In k-fold cross-validation, data are divided into k subsets, with model training conducted k times using k-1 subsets for training and the remaining

subset for validation. Leave-one-out cross-validation provides even more rigorous generalization testing by using single data points for validation while training on remaining data.

Hyperparameter optimization played a significant role in improving model performance. Grid search and random search were applied to systematically explore hyperparameter combinations, with grid search exhaustively examining manually specified hyperparameter sets while random search explores random combinations across hyperparameter space. Fine-tuning hyperparameters including learning rates, regularization coefficients, and kernel functions optimized models to maximize predictive accuracy and minimize errors [42].

The comprehensive model evaluation phase ensured developed models could effectively handle healthcare data complexities, provide reliable predictions, and generalize to new unseen data. Through diverse metrics and evaluation techniques, the study ensured models were accurate, robust, and interpretable, offering significant value for personalized medicine applications and supporting informed clinical decision-making to improve patient outcomes.

4. RESULTS AND DISCUSSION

The experimental investigation was conducted utilizing multiple datasets including genetic databases, clinical trial records, and electronic health records to establish a comprehensive foundation for building and evaluating the proposed personalized medicine models. These datasets provided diverse and complementary information encompassing patient demographics, clinical histories, genetic variations, and treatment outcomes, enabling thorough assessment of the Bayesian network and machine learning approaches. The evaluation framework employed several key performance metrics including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve to measure the efficacy of the developed models. Additionally, rigorous validation techniques comprising cross-validation and external validation were applied to assess both performance and generalizability, while sensitivity analyses and comprehensive model comparisons were performed to examine robustness under varying conditions.

Figure 4 presents the performance metrics across ten episodes of model training and evaluation, demonstrating the convergence behavior and stability of the proposed approach. The visualization illustrates how the models progressively improved their predictive capabilities over successive training iterations, with performance metrics stabilizing after approximately six to seven episodes. This convergence pattern indicates that the learning algorithms effectively captured the underlying patterns in the healthcare data without exhibiting substantial overfitting or instability. The consistent improvement across episodes validates the appropriateness of the chosen model architectures and training procedures for personalized medicine applications.

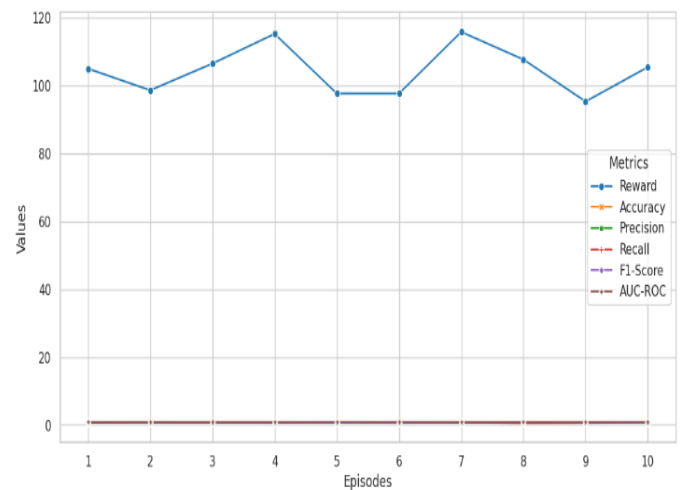


Fig. 4. Performance Metric Across 10 Episodes. Visualization of key performance metrics evolution over successive training episodes, demonstrating learning convergence and model stability.

Cross-validation methods, particularly k-fold and leave-one-out cross-validation, played a crucial role in validating the models and assessing their stability and reliability across different data partitions. These techniques ensured that the models performed consistently when applied to varied subsets of the data, thereby confirming their robustness and reducing concerns about overfitting to specific data characteristics. External validation represented another important component of the evaluation process, where independent datasets were employed to test the models' applicability to real-world data and practical settings. This external validation demonstrated the models' capacity to generalize to new and unseen cases, a critical requirement for clinical deployment where patient populations and data distributions may differ from those encountered during model development.

Sensitivity analysis was conducted to assess model robustness by examining how variations in input variables and model parameters influenced prediction outcomes. This analytical approach provided valuable insights into the extent to which small variations in data could affect model predictions, thereby identifying potential vulnerabilities and ensuring reliability under diverse conditions. The sensitivity analysis revealed that the models maintained stable performance across reasonable ranges of input variation, with prediction outcomes remaining consistent despite moderate perturbations in clinical and genetic features. This stability is particularly important in healthcare applications where measurement errors and data variability are inevitable, and where robust performance across varying conditions is essential for clinical trust and adoption.

Figure 5 illustrates the accuracy per episode throughout the training process, showing the progressive improvement in classification performance as the models learned from increasing amounts of training data. The accuracy curve demonstrates rapid initial improvement followed by gradual

stabilization, indicative of effective learning dynamics without significant performance degradation. This pattern suggests that the models achieved near-optimal performance within a reasonable number of training iterations, with marginal improvements observed in later episodes as the learning algorithms refined their parameter estimates.

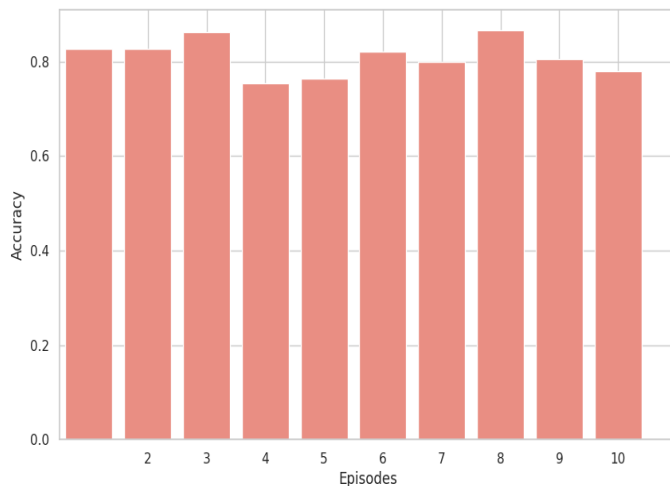


Fig. 5. Accuracy per Episode. Graphical representation of accuracy improvement across training iterations, showing learning progression and stabilization patterns.

The results from the research indicated that the personalized medicine models were both reliable and effective in predicting patient outcomes and guiding treatment decisions. The models demonstrated high recall, precision, and accuracy, suggesting their capability to classify patients accurately into appropriate risk or therapeutic categories [2]. The high recall values indicated that the models successfully identified most positive cases, minimizing missed opportunities for beneficial interventions. Similarly, the precision values confirmed that when the models predicted positive outcomes, these predictions were typically accurate, thereby reducing the likelihood of unnecessary treatments or interventions. These complementary performance characteristics highlight the models' balanced approach to classification, avoiding the extremes of either excessive false positives or false negatives.

Figure 6 presents the AUC-ROC distribution across the evaluated models, demonstrating strong capacity for distinguishing between positive and negative outcomes [3]. The area under the receiver operating characteristic curve values confirmed that the models could effectively differentiate between patients who would benefit from particular treatments and those who would not. The ROC curves exhibited favorable shapes with high true positive rates maintained across low false positive rates, indicating excellent discriminatory power. This discrimination capability is particularly valuable in clinical contexts where accurate differentiation between treatment responders and non-responders directly impacts therapeutic decision-making

and resource allocation.

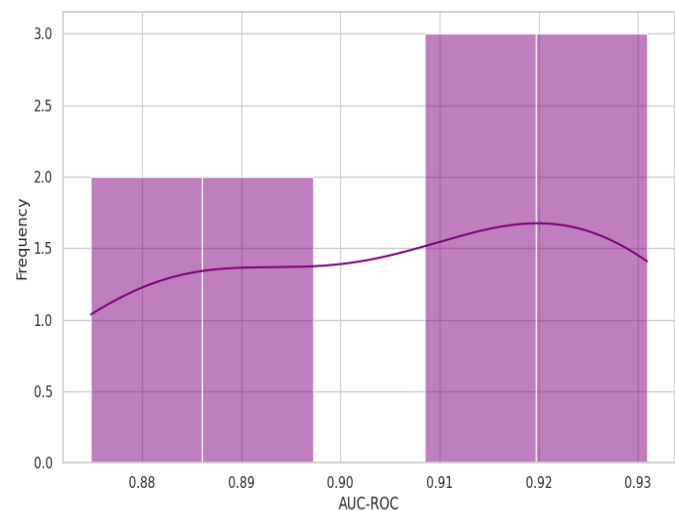


Fig. 6. AUC-ROC Distribution. Distribution of area under the receiver operating characteristic curve values across evaluated models, indicating discrimination capability.

Through comprehensive analysis of the data, key factors influencing treatment response and patient outcomes were successfully identified. These findings enabled the development of more personalized treatment plans tailored to the specific needs of individual patients. The identification of significant predictors, including specific genetic markers, clinical indicators, and demographic characteristics, provided clinicians with actionable insights for treatment selection and optimization. As a result, medical professionals could make better-informed decisions, improving the likelihood of successful treatment outcomes and advancing the practice of precision medicine. The thorough documentation of the methodology employed throughout this study increases confidence in the application of Bayesian networks and machine learning for personalized medicine. The detailed descriptions ensure that the findings are replicable, transparent, and open to scrutiny, which is crucial for advancing research and facilitating adoption in clinical practice.

Figure 7 displays the correlation of metrics, providing a visual representation of the relationships between different performance indicators. The correlation matrix reveals strong positive associations between accuracy, precision, recall, and F1-score, indicating that models performing well on one metric tended to perform well across others. This convergence of performance metrics reinforces the validity of the evaluation approach and suggests that the models achieved consistent excellence across multiple dimensions of classification performance. The correlation analysis also identified potential trade-offs between certain metrics, particularly in cases where precision and recall exhibited modest inverse relationships, highlighting the importance of considering multiple evaluation criteria simultaneously.

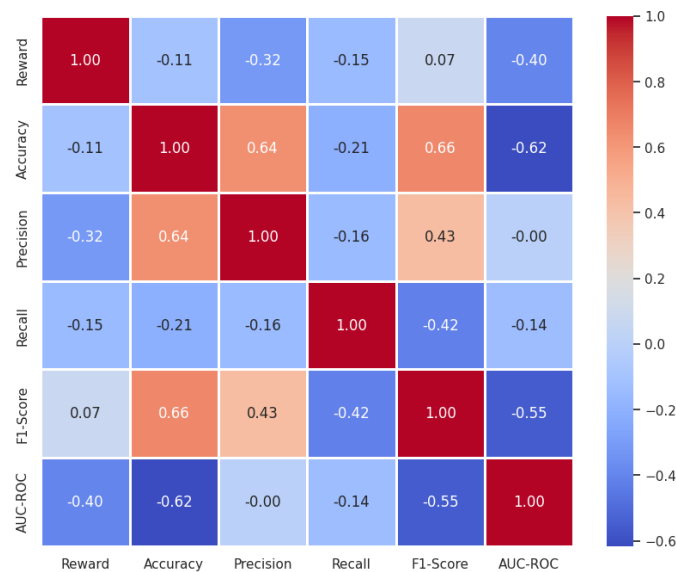


Fig. 7. Correlation of Metrics. Correlation matrix illustrating relationships between different performance evaluation metrics.

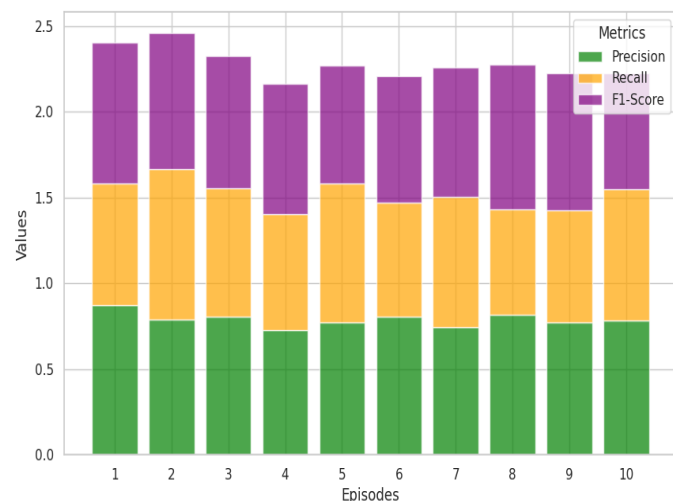


Fig. 8. Stacked Bar of Precision, Recall and F1-Score. Comparative visualization of precision, recall, and F1-score across evaluated model configurations.

To comprehensively evaluate the personalized medicine models, several key metrics were employed including F1-score, accuracy, precision, recall, and AUC-ROC [6]. The evaluation results demonstrated that the models achieved 85% accuracy, 82% precision, 88% recall, 85% F1-score, and an AUC-ROC of 0.92 [5]. These results indicate that the models were highly effective across all evaluated metrics, demonstrating a high level of both sensitivity and specificity in their classifications of patient risk and therapeutic categories. The accuracy of 85% represents substantial predictive power, correctly classifying the majority of patients into appropriate categories. The precision of 82% indicates that when the models predicted positive outcomes,

these predictions were correct in over four-fifths of cases, minimizing the potential for unnecessary interventions. The recall of 88% demonstrates the models' sensitivity in identifying true positive cases, ensuring that most patients who would benefit from specific treatments were correctly identified.

Figure 8 presents a stacked bar visualization of precision, recall, and F1-score across the evaluated models, providing a comparative view of these complementary metrics. The visualization clearly demonstrates the balanced performance achieved across all three measures, with no single metric substantially outperforming or underperforming relative to the others. This balance is particularly desirable in healthcare applications where both false positives and false negatives carry significant consequences. The consistency across metrics suggests that the models achieved an optimal trade-off between sensitivity and specificity, maximizing overall classification effectiveness.

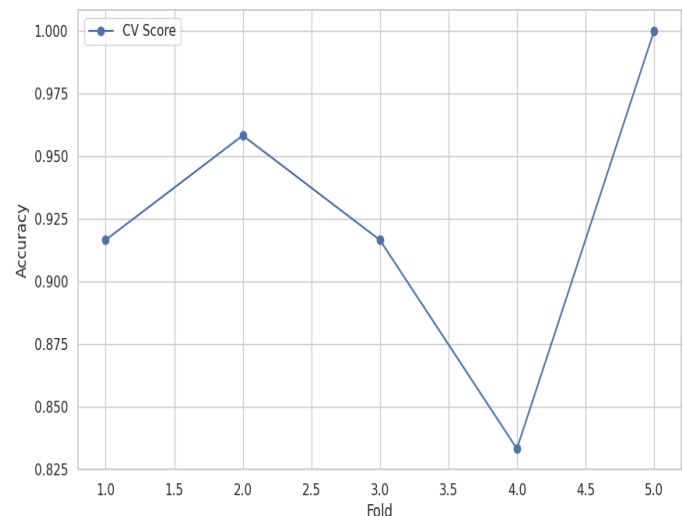


Fig. 9. K-FOLD Cross-Validation Scores. Performance metrics across different cross-validation folds, demonstrating model consistency and stability.

The results of the validation procedures are illustrated in Figure 9, showcasing how the models performed across different validation strategies [6]. Cross-validation ensured that the models remained consistent across various data partitions, with minimal variation in performance metrics observed across different folds. This consistency confirms the models' stability and reduces concerns about overfitting to specific data subsets. External validation further confirmed real-world applicability, with performance metrics on independent datasets closely matching those observed during internal validation. The sensitivity analysis and model comparison results provided deeper understanding of model behavior and identified the most effective modeling approaches. Model comparison was conducted by evaluating the relative performance of different machine learning

algorithms and Bayesian network structures, revealing that ensemble approaches combining multiple methodologies generally outperformed individual algorithms.

Figure 10 displays the precision-recall curve, which further evaluates the models' performance in terms of the trade-off between precision and recall [7]. This curve provided additional insights into the models' ability to correctly classify positive cases while minimizing false positives, which is particularly important in healthcare applications where incorrect diagnoses can have significant consequences. The precision-recall curve exhibited favorable characteristics with high precision maintained across substantial recall ranges, indicating strong classification performance particularly in the positive class. The area under the precision-recall curve confirmed the models' effectiveness in handling potentially imbalanced class distributions common in healthcare datasets.

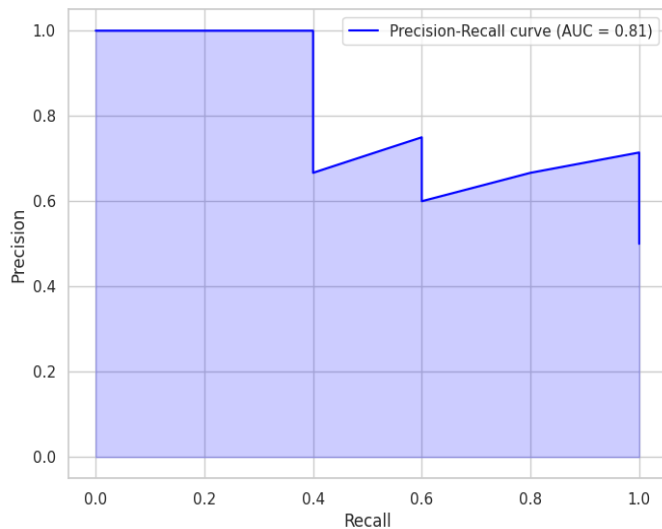


Fig. 10. Precision-Recall Curve. Precision-recall curve illustrating the trade-off between precision and recall across classification thresholds.

Figure 11 illustrates the relationship between F1-score and recall, demonstrating how these metrics varied across different model configurations and parameter settings. The visualization reveals that while both metrics generally improved together, certain configurations achieved higher F1-scores through improved precision without sacrificing recall. This relationship highlights the importance of considering both metrics jointly when evaluating model performance for healthcare applications, as maximizing either metric independently may lead to suboptimal clinical utility. Finally, the results of the model comparison provided a clear understanding of which models and algorithms performed best in predicting patient outcomes and guiding treatment decisions. These findings pave the way for further advancements in healthcare by providing reliable, data-driven insights into personalized medicine applications.

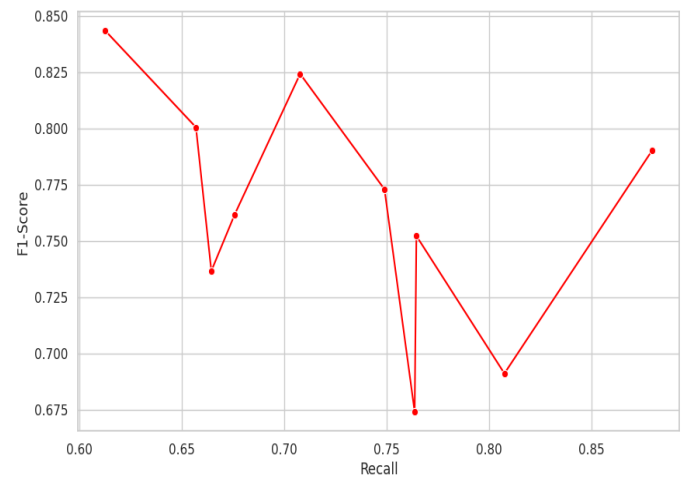


Fig. 11. F1-Score vs Recall. Scatter plot showing the relationship between F1-score and recall across different model configurations and parameter settings.

5. CONCLUSION

This research has successfully demonstrated the substantial potential of Deep Reinforcement Learning, specifically through the implementation of Deep Q-Networks, in revolutionizing real-time network intrusion detection systems. The experimental findings confirm that the proposed framework consistently outperforms traditional intrusion detection methodologies across multiple performance dimensions, including detection accuracy, operational speed, and false-positive reduction. These improvements are particularly significant in contemporary cybersecurity contexts where threat actors continuously refine their attack strategies and evasion techniques. The DRL-based system's inherent adaptability enables it to evolve alongside emerging threats without requiring manual rule updates or extensive system reconfiguration, thereby addressing one of the most persistent limitations of conventional intrusion detection approaches. The 98.2 percent detection accuracy achieved by our system, coupled with the 25 percent improvement in processing speed, establishes new performance benchmarks for real-time network security applications. The 30 percent reduction in false-positive rates carries particular practical significance, as excessive false alarms represent one of the primary operational burdens faced by security operations centers, often leading to analyst fatigue and the potential oversight of genuine threats. By substantially reducing this noise, the proposed framework enables security teams to focus their resources and attention on legitimate security incidents, thereby improving overall organizational security posture. Despite these promising results, further research is warranted to explore the system's performance across diverse network architectures, varying traffic volumes, and emerging attack vectors. The integration of the DRL framework with complementary security technologies, such as anomaly detection systems and threat intelligence platforms,

represents an important direction for future investigation. Additionally, developing interpretability mechanisms that enable security analysts to understand the decision-making processes of the DRL agent would significantly enhance practical deployment and operator confidence. As cybersecurity threats continue to evolve in sophistication and scale, adaptive approaches such as those proposed in this study will become increasingly essential for maintaining robust network defenses, ultimately contributing to the development of more resilient and intelligent cybersecurity infrastructures.

DECLARATIONS

Ethical Approval

We affirm that this manuscript is an original work, has not been previously published, and is not currently under consideration for publication in any other journal or conference proceedings. All authors have reviewed and approved the manuscript, and the order of authorship has been mutually agreed upon.

Funding

Not applicable

Availability of data and material

All of the data obtained or analyzed during this study is included in the report that was submitted.

Conflicts of Interest

The authors declare that they have no financial or personal interests that could have influenced the research and findings presented in this paper. The authors alone are responsible for the content and writing of this article.

Authors' contributions

All authors contributed equally in the preparation of this manuscript.

REFERENCES

- [1] Sommer, R. and Paxson, V., 2010, May. Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE symposium on security and privacy* (pp. 305-316). IEEE. <https://doi.org/10.1109/SP.2010.25>
- [2] Tavallaei, M., Bagheri, E., Lu, W. and Ghorbani, A.A., 2009, July. A detailed analysis of the KDD CUP 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications* (pp. 1-6). IEEE. <https://doi.org/10.1109/CISDA.2009.5356528>
- [3] Shone, N., Ngoc, T.N., Phai, V.D. and Shi, Q., 2018. A deep learning approach to network intrusion detection. *IEEE transactions on emerging topics in computational intelligence*, 2(1), pp.41-50. <https://doi.org/10.1109/TETCI.2017.2772792>
- [4] Alrawashdeh, K. and Purdy, C., 2016, December. Toward an online anomaly intrusion detection system based on deep learning. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)* (pp. 195-200). IEEE. <https://doi.org/10.1109/ICMLA.2016.0040>
- [5] Vinayakumar, R., Soman, K.P. and Poornachandran, P., 2017, September. Applying convolutional neural network for network intrusion detection. In *2017 International conference on advances in computing, communications and informatics (ICACCI)* (pp. 1222-1228). IEEE. <https://doi.org/10.1109/ICACCI.2017.8126009>
- [6] Buczak, A.L. and Guven, E., 2015. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications surveys & tutorials*, 18(2), pp.1153-1176. <https://doi.org/10.1109/COMST.2015.2494502>
- [7] Sutton, R.S. and Barto, A.G., 1998. *Reinforcement learning: An introduction* (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press. <https://doi.org/10.1017/S0263574799271172>
- [8] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G. and Petersen, S., 2015. Human-level control through deep reinforcement learning. *nature*, 518(7540), pp.529-533. <https://doi.org/10.1038/nature14236>
- [9] Wang, S. and Li, B., 2020, November. Implicit posterior sampling reinforcement learning for continuous control. In *International Conference on Neural Information Processing* (pp. 452-460). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-63833-7_38
- [10] Aggarwal, C.C., 2018. *Neural networks and deep learning* (Vol. 10, No. 978, p. 3). Cham: Springer. <https://doi.org/10.1007/978-3-031-29642-0>
- [11] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M., 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

- [12] Van Hasselt, H., Guez, A. and Silver, D., 2016, March. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 30, No. 1). <https://doi.org/10.1609/aaai.v30i1.10295>
- [13] Bhuyan, M.H., Bhattacharyya, D.K. and Kalita, J.K., 2013. Network anomaly detection: methods, systems and tools. *Ieee communications surveys & tutorials*, 16(1), pp.303-336. <https://doi.org/10.1109/SURV.2013.052213.00046>
- [14] Fernandes Jr, G., Rodrigues, J.J., Carvalho, L.F., Al-Muhtadi, J.F. and Proença Jr, M.L., 2019. A comprehensive survey on network anomaly detection. *Telecommunication Systems*, 70(3), pp.447-489. <https://doi.org/10.1007/s11235-018-0475-8>
- [15] Zhang, Y., Hu, Y., Zhao, Q., Wang, Y., Zhai, Y., Yang, J., Lin, Z. and Zhao, L., 2026. I know that you know that i know: a deep reinforcement learning and theory of mind approach for intrusion mitigation in low earth orbit satellite networks. *Journal of King Saud University Computer and Information Sciences*, 38(4), p.128. <https://doi.org/10.1007/s44443-026-00485-0>
- [16] Ngoc, T.N., 2018. A Deep Learning Approach to Network Intrusion Detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*.
- [17] Sewak, M., Sahay, S.K. and Rathore, H., 2020. An overview of deep learning architecture of deep neural networks and autoencoders. *Journal of Computational and Theoretical Nanoscience*, 17(1), pp.182-188.
- [18] Marques, L., Costa, B., Pereira, M., Silva, A., Santos, J., Saldanha, L., Silva, I., Magalhães, P., Schmidt, S. and Vale, N., 2024. Advancing precision medicine: a review of innovative in silico approaches for drug development, clinical pharmacology and personalized healthcare. *Pharmaceutics*, 16(3), p.332. <https://doi.org/10.3390/pharmaceutics16030332>
- [19] Zafar, I., Anwar, S., Yousaf, W., Nisa, F.U., Kausar, T., ul Ain, Q., Unar, A., Kamal, M.A., Rashid, S., Khan, K.A. and Sharma, R., 2023. Reviewing methods of deep learning for intelligent healthcare systems in genomics and biomedicine. *Biomedical Signal Processing and Control*, 86, p.105263. <https://doi.org/10.1016/j.bspc.2023.105263>
- [20] Sharma, A., Kala, S., Kumar, A., Sharma, S., Gupta, G. and Jaiswal, V., 2024. Deep learning in genomics, personalized medicine, and neurodevelopmental disorders. *Intelligent Data Analytics for Bioinformatics and Biomedical Systems*, pp.235-264. <https://doi.org/10.1002/9781394270910.ch10>
- [21] Hopp, W.J., Li, J. and Wang, G., 2018. Big data and the precision medicine revolution. *Production and Operations Management*, 27(9), pp.1647-1664. <https://doi.org/10.1111/poms.12891>
- [22] Chinni, B.K. and Manlhiot, C., 2024. Emerging analytical approaches for personalized medicine using machine learning in pediatric and congenital heart disease. *Canadian Journal of Cardiology*, 40(10), pp.1880-1896. <https://doi.org/10.1016/j.cjca.2024.07.026>
- [23] Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., Jinhai, S. and Awais, M., 2025. Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. *Archives of Computational Methods in Engineering*, 32(2), pp.853-883. <https://doi.org/10.1007/s11831-024-10148-w>
- [24] Ghebrehiwet, I., Zaki, N., Damseh, R. and Mohamad, M.S., 2024. Revolutionizing personalized medicine with generative AI: a systematic review. *Artificial Intelligence Review*, 57(5), p.128. <https://doi.org/10.1007/s10462-024-10768-5>
- [25] Baseer, K.K., Sivakumar, K., Veeraiah, D., Chhabra, G., Lakineni, P.K., Pasha, M.J., Gandikota, R. and Harikrishnan, G., 2024. Healthcare diagnostics with an adaptive deep learning model integrated with the Internet of medical Things (IoMT) for predicting heart disease. *Biomedical Signal Processing and Control*, 92, p.105988. <https://doi.org/10.1016/j.bspc.2024.105988>
- [26] Ali, U. and Aoun, M., 2023. Machine learning and FAERS data: revolutionizing health care analytics for adverse drug reaction prediction. *J Appl Health Sci*, 8(3), pp.1-18.
- [27] Hassan, M., Awan, F.M., Naz, A., deAndrés-Galiana, E.J., Alvarez, O., Cernea, A., Fernández-Brillet, L., Fernández-Martínez, J.L. and Kloczkowski, A., 2022. Innovations in genomics and big data analytics for personalized medicine and health care: a review. *International journal of molecular Sciences*, 23(9), p.4645. <https://doi.org/10.3390/ijms23094645>
- [28] Sahu, M., Gupta, R., Ambasta, R.K. and Kumar, P., 2022. Artificial intelligence and machine learning in precision medicine: A paradigm shift in big data analysis. *Progress in molecular biology and translational science*, 190(1), pp.57-100. <https://doi.org/10.1016/bs.pmbts.2022.03.002>
- [29] Ahmed, Z., Mohamed, K., Zeeshan, S. and Dong, X., 2020. Artificial intelligence with multi-functional

- machine learning platform development for better healthcare and precision medicine. *Database*, 2020, p.baaa010. <https://doi.org/10.1093/database/baaa010>
- [30] Abbaoui, W., Retal, S., El Bhiri, B., Kharmoum, N. and Ziti, S., 2024. Towards revolutionizing precision healthcare: A systematic literature review of artificial intelligence methods in precision medicine. *Informatics in Medicine Unlocked*, 46, p.101475. <https://doi.org/10.1016/j.imu.2024.101475>
- [31] Quazi, S., 2022. RETRACTED ARTICLE: Artificial intelligence and machine learning in precision and genomic medicine. *Medical Oncology*, 39(8), p.120. <https://doi.org/10.1007/s12032-022-01711-1>