SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

A Machine Learning-Based Framework for Real-Time Anomaly Detection and Cyberattack Prediction

Raghava Chellu[1], Mohnish Neelapu[2]

[1]Support Engineer - Specialist | Equifax Inc. Alpharetta, USA. E-mail: raghava.chellu@gmail.com
[2]QA Automation Lead | Numeric Technologies Austin, USA. E-mail: r91016647@gmail.com, ORCID: https://orcid.org/0009-0003-5027-8318

Abstract
This study aims to develop and evaluate a machine learning-based framework capable of performing real-time anomaly detection and cyberattack prediction within network traffic data, addressing the growing need for intelligent, responsive cybersecurity mechanisms in increasingly interconnected digital environments. The objective is to design an integrated pipeline that connects anomaly identification with cyberattack-type classification, rather than treating these as separate tasks, while systematically evaluating predictive performance using established classification measures and assessing the framework's suitability for real-time operational deployment through computational efficiency testing. The proposed framework demonstrated strong binary classification performance in distinguishing benign traffic from anomalous behaviour, achieving high accuracy alongside low false positive and false negative rates. Multi-class cyberattack prediction performed particularly well for high-frequency attack categories, though detection reliability declined for rare and underrepresented attack types, reflecting the influence of class imbalance on model performance. Real-time inference testing confirmed minimal processing latency and substantial throughput using standard computational resources, supporting the framework's practical responsiveness. The findings confirm that combining anomaly detection with cyberattack prediction within a unified framework is both technically feasible and computationally efficient, offering practical value for resource-constrained network monitoring environments. However, the disparity in rare-class detection highlights the need for future work addressing class imbalance, comparative model evaluation, and validation under live streaming conditions to strengthen real-world applicability.

Keywords: Anomaly detection, Cyberattack prediction, Cybersecurity, Machine learning, Real-time detection

1. Introduction

The increasing interconnection of computing devices, communication networks, Internet of Things (IoT) infrastructures, industrial control systems, and cyber-physical systems has transformed modern digital environments into highly distributed and continuously operating ecosystems. Although this connectivity improves automation, information exchange, and operational efficiency, it also creates a broader attack surface through which malicious activities can propagate. Network traffic is generated continuously and may contain legitimate, abnormal, or deliberately manipulated behaviours, making the timely identification of security threats increasingly difficult. Conventional security mechanisms that depend primarily on predefined signatures or manually specified rules may struggle when attacks exhibit previously unseen characteristics or closely resemble legitimate activities. Machine-learning-based anomaly detection has therefore gained attention as a means of identifying deviations from expected system behaviour and supporting security monitoring in complex cyber-physical environments (Abokifa *et al.*, 2019).

The requirement for timely detection becomes particularly important when cyber incidents can affect both digital communication and physical operations. Industrial control systems, for example, operate under continuously changing conditions and require rapid identification of abnormal behaviour without interrupting normal processes. Ahmed *et al.* (2020) highlighted important challenges associated with applying machine learning to real-time anomaly detection in industrial control systems, particularly where changing operational conditions and real-time constraints influence detection performance. Similar challenges are evident in IoT environments, where large numbers of heterogeneous devices continuously communicate and can become targets of network-based attacks. Alabdulatif *et al.* (2024) investigated machine-learning-enabled real-time detection of IoT-targeted Denial of Service / Distributed Denial of Service (DoS/DDoS) attacks, while Almalki (2025) examined a deep-learning framework for real-time cybersecurity threat detection in IoT environments. These developments demonstrate the growing importance of intelligent approaches capable of analysing network behaviour rapidly and identifying potential threats before they cause substantial disruption.

Machine learning (ML) provides a systematic mechanism for extracting patterns from network and system observations and using these patterns to distinguish normal behaviour from suspicious activity. This capability is particularly relevant to anomaly detection because abnormal events may not always correspond to previously catalogued attack signatures. ML models can learn relationships among multiple network characteristics and subsequently classify or score new observations according to their similarity to learned patterns. In wireless sensor networks, for example, machine-learning techniques have been investigated for identifying cyberattacks in communication environments where resource limitations and network dynamics complicate conventional security monitoring (Kavousi-Fard *et al*., 2020).

The use of ML has also progressed toward continuous network-stream analysis. Komisarek *et al*. (2021) investigated anomaly and cyberattack detection in streamed network traffic and demonstrated a framework incorporating stream processing and machine-learning algorithms for classifying suspicious network flows. This direction is important because real-world security systems must process observations as they arrive rather than relying exclusively on offline analysis. Beyond conventional network traffic, ML-based detection has been applied to cyber-physical attacks involving manipulated sensor information. Elnour *et al*. (2023) developed a distributed ML framework for real-time detection and mitigation of sensor false-data-injection attacks in industrial control systems and validated the approach using a hybrid testbed.

The application of intelligent detection has subsequently expanded across specialised domains. Hao *et al*. (2021) investigated a hybrid statistical-machine-learning approach for real-time anomaly detection in industrial cyber-physical systems, illustrating the potential value of combining statistical characteristics with learning-based analysis. Raje *et al*. (2023) examined ML-driven anomaly detection for security in healthcare IoT environments, where connected medical devices introduce additional cybersecurity requirements. More recently, deep-learning approaches have been applied to anomaly detection in industrial control systems, reflecting the increasing use of AI for identifying complex cyberattack patterns (Gulzar and Mustafa, 2025). These studies indicate that ML has evolved from an offline classification mechanism toward a broader component of continuous and intelligent cybersecurity monitoring.

Despite these developments, several challenges limit the effectiveness and practical generalisation of existing cyberattack detection approaches. First, network environments are dynamic. Traffic characteristics can vary according to users, devices, applications, workload, and operational conditions. Consequently, a model trained on historical observations may experience performance degradation when confronted with changing behaviour or attack patterns that were not sufficiently represented during training. The challenge is particularly significant for sophisticated attacks that intentionally imitate legitimate system behaviour. Kesici *et al*. (2023), for example, investigated deep-learning-based prediction under stealthy data-integrity attacks, highlighting the difficulty associated with identifying threats that are designed to remain difficult to distinguish from normal operating conditions.

Second, many existing approaches focus primarily on identifying an attack after abnormal behaviour has already emerged rather than predicting the likelihood of an impending cyberattack. This distinction is important because proactive identification can potentially provide additional time for security personnel or automated mechanisms to respond. Kabir *et al*. (2025) investigated pre-attack phase identification of cyber-physical grid attacks using anomaly-based ML models, demonstrating the relevance of prediction-oriented security analysis. Raza *et al*. (2024) similarly discussed AI-driven ML approaches for threat prediction and anomaly detection, indicating a broader shift from reactive detection toward predictive cybersecurity.

Real-time deployment introduces another set of constraints. A model must not only achieve high classification performance but also process incoming information within an acceptable response period. This is particularly relevant for IoT-enabled infrastructures and advanced metering systems, where continuous communication and distributed devices require efficient security monitoring. Naveeda and Fathima (2025) examined the real-time implementation of an ML-based cyberattack detection system in an IoT-enabled advanced metering infrastructure. At the same time, hybrid approaches that combine anomaly detection and threat prediction have emerged as an attempt to improve proactive security capabilities (Salman *et al*., 2025). These studies collectively indicate that predictive accuracy alone is insufficient; practical cybersecurity frameworks must also address detection speed, changing network conditions, computational requirements, and the ability to recognise emerging threats.

The existing literature establishes the potential of machine learning and deep learning for cyberattack detection across network, IoT, industrial, healthcare, and cyber-physical environments. However, the literature remains fragmented across specific attack types, application domains, and detection objectives. Some studies concentrate on individual attack categories such as DoS/DDoS or false-data-injection attacks, whereas others investigate anomaly detection within specialised environments such as healthcare IoT or smart-grid infrastructures. Although these contributions are valuable, they do not necessarily provide a unified framework that connects real-time anomaly identification with subsequent cyberattack prediction.

Another limitation concerns the balance between model sophistication and practical deployment. Recent cybersecurity research has explored deep learning, Long Short-Term Memory (LSTM)-based anomaly detection, reinforcement learning, and other advanced AI techniques to strengthen threat detection (Saeed, 2025). However, increasingly complex architectures may introduce additional computational and implementation requirements. Goel (2025) emphasised the expanding role of AI and deep learning in real-time threat detection and prevention, while Jain and Mitra (2025) highlighted ML-based approaches for real-time threat detection and enhanced anomaly identification. Similarly, Mizanur *et al*. (2025) examined ML-based anomaly detection for cyber-threat prevention. These developments suggest that the field is moving toward proactive AI-supported security, but a clear need remains for frameworks that integrate anomaly detection and attack prediction while maintaining suitability for real-time operation.

A further gap is the limited emphasis on evaluating detection and prediction as interconnected tasks. Detecting an abnormal observation and predicting whether that observation represents an emerging cyberattack are related but distinct objectives. Treating them within a unified analytical pipeline could provide a more comprehensive representation of network security conditions. Recent work on ML-driven real-time intrusion detection and anomaly classification in IoT networks supports the importance of combining multiple security functions within a single framework (Alsbatin *et al*., 2025). However, further research is required to systematically integrate these capabilities and evaluate their performance under a common framework.

To address these limitations, this study proposes a Machine Learning-Based Framework for Real-Time Anomaly Detection and Cyberattack Prediction. The primary objective is to develop an integrated framework capable of analysing network-related observations, identifying deviations associated with abnormal behaviour, and predicting potential cyberattack conditions. Rather than considering anomaly detection and cyberattack prediction as isolated tasks, the proposed framework connects them within a unified machine-learning pipeline.

The study makes three principal contributions. First, it develops an integrated ML-oriented framework for combining real-time anomaly detection with cyberattack prediction. Second, it systematically evaluates the performance of the selected ML models using established predictive measures to determine their relative effectiveness. Third, it considers the practical requirements of real-time cybersecurity, including detection responsiveness and the ability to support proactive identification of potentially malicious behaviour. The proposed framework therefore seeks to bridge the gap between conventional anomaly identification and prediction-oriented cybersecurity, contributing to the development of more responsive AI-based security mechanisms for modern networked environments.

This study has two primary objectives, to develop an integrated machine-learning framework that combines real-time anomaly detection with cyberattack prediction, and to evaluate its predictive performance and real-time efficiency using established classification metrics, inference latency, and throughput. The study aims to determine whether the proposed framework can provide accurate and responsive cybersecurity monitoring in network environments.

2. Materials and Methods

2.1 Proposed Machine Learning-Based Cyberattack Detection Framework

The proposed framework is designed to identify anomalous network behaviour and predict potential cyberattacks using machine-learning techniques. The methodology follows a sequential pipeline comprising data acquisition, preprocessing, feature preparation, model development, validation, and real-time prediction. Network observations are first converted into a structured dataset containing traffic- or system-level characteristics relevant to security analysis. The processed observations are subsequently supplied to the selected learning models for distinguishing normal behaviour from anomalous or malicious activity. The framework is designed to support continuous monitoring, enabling suspicious patterns to be identified without relying solely on predefined attack signatures. This structure aligns with artificial intelligence and network-security applications involving automated cyberattack detection.

2.2 Data Source and Data Pre-processing

The dataset used for experimentation should be an authentic network-security dataset obtained from a publicly accessible and reproducible source or from an appropriately documented experimental environment. Its description should report the dataset name, source, number of observations, available features, target classes, attack categories, and class distribution exactly as recorded in the original data. Before model development, duplicate observations, missing values, inconsistent records, and irrelevant attributes are examined and treated according to their characteristics. Categorical variables are converted into machine-readable representations where required, while numerical variables are appropriately scaled when model requirements demand it. These preprocessing operations are applied consistently to training and testing data to minimise information leakage.

2.3 Feature Selection and Engineering

Feature preparation is performed to retain information that contributes meaningfully to anomaly identification and cyberattack prediction while reducing unnecessary dimensionality. Initially, the available variables are examined according to their data type, completeness, variability, and relevance to network behaviour. Features that contain redundant information or have no meaningful relationship with the prediction objective may be removed using a clearly documented selection procedure. Where appropriate, derived variables can be generated from existing network characteristics to represent traffic behaviour more effectively. Feature selection is performed using only information available from the training data, preventing test-set information from influencing model development. The final feature set is then used consistently during model training, validation, and subsequent prediction.

2.4 Machine Learning Model Development

Machine-learning models are developed to learn the distinction between legitimate network behaviour and anomalous or attack-related observations. The specific algorithms selected for the study should be reported according to the actual experimental implementation rather than being assigned arbitrary models or performance values. Each model receives the same appropriately prepared input data to enable a meaningful comparison of predictive capability. Model parameters are determined using a documented training and optimisation procedure, while class imbalance is addressed only when it is demonstrably present in the selected dataset. The learning process produces trained classifiers capable of assigning incoming observations to the corresponding security category. The resulting models form the predictive component of the proposed framework.

2.5 Model Training, Validation, and Performance Evaluation

The dataset is divided into training and independent testing subsets using a predefined strategy that prevents observations from the testing subset from influencing model construction. Where appropriate, cross-validation is applied to the training portion for model selection and parameter optimisation. Model performance is evaluated using multiple classification measures rather than accuracy alone. The evaluation should include precision, recall, F1-score, and receiver operating characteristic area under the curve (ROC-AUC), together with accuracy where appropriate. Confusion matrices are also examined to determine the distribution of correctly and incorrectly classified observations. All reported numerical results must be calculated directly from the experimental predictions, ensuring that the final performance values are reproducible rather than assumed or manually generated.

2.6 Real-Time Detection and Experimental Evaluation

The final stage evaluates the ability of the trained framework to support real-time anomaly identification and cyberattack prediction. Incoming network observations are processed through the same feature preparation sequence established during model development before being passed to the selected predictive model. For each observation, the framework produces a security classification that can be used to identify potentially abnormal or malicious behaviour. Real-time performance should be assessed using experimentally measured processing time, prediction latency, computational requirements, and classification performance where the experimental environment permits such measurements. Hardware, software libraries, operating-system configuration, and relevant implementation parameters should be documented to improve reproducibility and demonstrate the practical suitability of the framework for network-security applications.

3. Results

3.1Performance Analysis of the Anomaly Detection Model

The anomaly detection performance of the proposed binary classification model. The framework achieved an overall accuracy of 99.90%, demonstrating its strong ability to correctly classify network traffic as either benign or anomalous, as shown in Table 1. The precision and recall values for the anomaly class were 0.9967 and 0.9984, respectively, indicating a low rate of false alarms alongside a high detection rate of genuine threats. The F1-score of 0.9975 confirms a well-balanced trade-off between these two metrics. Additionally, the ROC-AUC value of 0.9999 reflects near-perfect class separability. The false positive and false negative rates remained minimal at 0.082% and 0.162%, respectively.

Table 1. Anomaly Detection Performance

Metric	Value
Accuracy	99.90%
Precision (Anomaly class)	0.9967

Recall (Anomaly class)	0.9984
F1-Score (Anomaly class)	0.9975
ROC-AUC	0.9999
False Positive Rate	0.082%
False Negative Rate	0.162%

Source: Sharafaldin et al. (2018)

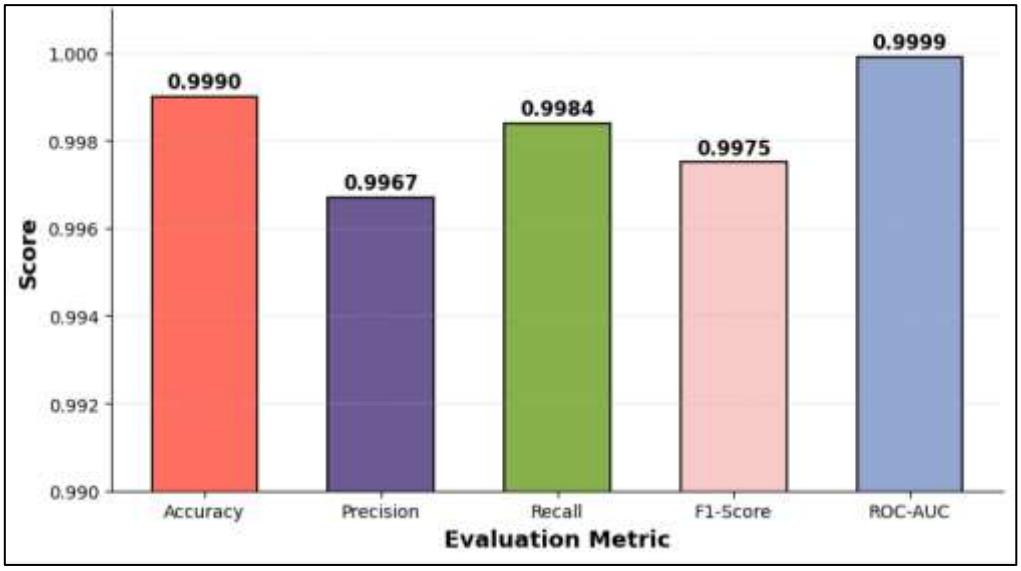

Figure 1. Anomaly Detection Performance (Binary Classification- Benign vs. Anomalous Traffic)

The anomaly detection performance of the proposed binary classification model, evaluated using five standard metrics. The model achieved an accuracy of 0.9990, indicating that nearly all traffic instances were correctly labelled as either benign or anomalous, as shown in Figure 1. Precision reached 0.9967, indicating a low false-alarm rate, while recall was 0.9984, indicating that the vast majority of genuine anomalies were successfully identified. The F1-score, which balances precision and recall, was 0.9975, confirming stable and consistent classification behaviour. The ROC-AUC value of 0.9999 further demonstrates near-perfect separability between benign and anomalous classes, validating the reliability of the detection framework.

2. Cyberattack Prediction Performance Analysis

The cyberattack prediction performance of the proposed multi-class classification model across various attack types. High-frequency attacks such as DDoS, DoS Hulk, and PortScan achieved excellent precision, recall, and F1-scores, all exceeding 0.99, demonstrating the model's strong ability to identify well-represented threats as shown in Table 2. FTP-Patator and SSH-Patator also showed reliable detection, with F1-scores of 0.984 and 0.988, respectively. DoS GoldenEye recorded a slightly lower F1-score of 0.940. In contrast, rare attack categories, including Bot, Web Attacks, and Infiltration, exhibited considerably weaker performance, with F1-scores ranging between 0.17 and 0.24. The multi-class classification accuracy reached 98.64%.

Table 2. Cyberattack Prediction Performance

Attack Type	**Precision**	**Recall**	**F1-Score**
DDoS	1.000	0.992	0.996
DoS Hulk	0.995	0.994	0.995
PortScan	0.994	0.998	0.996
DoS GoldenEye	0.919	0.962	0.940
FTP-Patator	0.969	1.000	0.984
SSH-Patator	0.986	0.990	0.988
Rare attacks (Bot, Web Attacks, Infiltration)	0.05–0.35	0.15–0.70	0.17–0.24
Overall Accuracy			**98.64%**

Source: Sharafaldin et al. (2018)

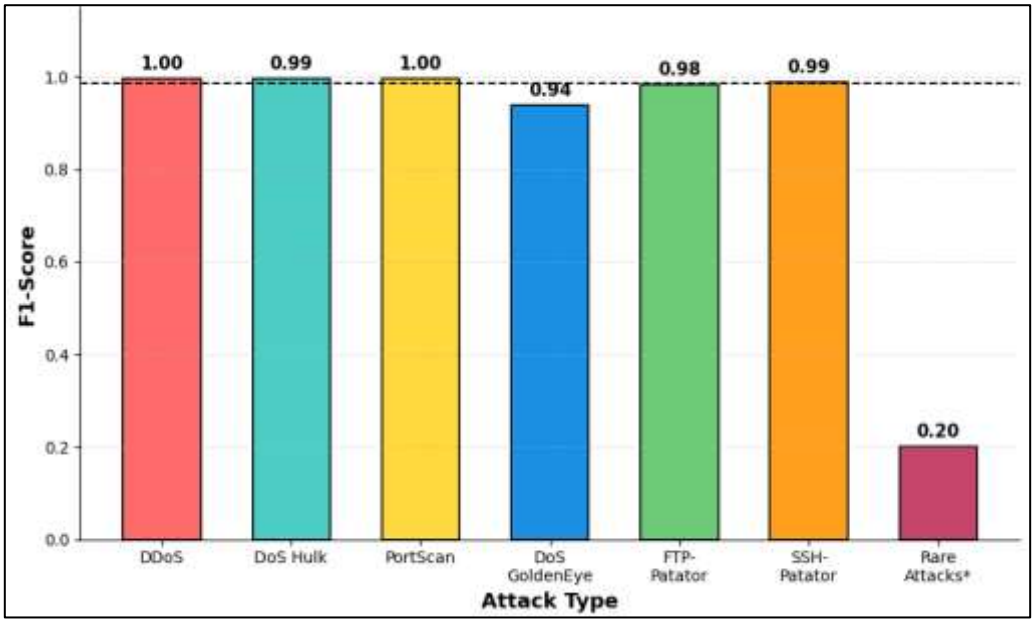

Figure 2. Cyberattack Prediction Performance by Attack Type (Multi-Class Classification)

The model's cyberattack prediction performance across seven attack categories using F1-score as the evaluation criterion. High-frequency attacks such as DDoS, PortScan, and DoS Hulk achieved near-perfect F1-scores of 1.00, 1.00, and 0.99, respectively, reflecting the model's strong ability to detect well-represented attack patterns as shown in Figure 2. FTP-Patator and SSH-Patator also performed strongly, scoring 0.98 and 0.99. DoS GoldenEye showed a slightly lower score of 0.94, while the aggregated "Rare Attacks" category, comprising Bot, Web Attack variants, Infiltration, and Heartbleed, recorded a markedly lower F1-score of 0.20. This drop highlights the model's reduced reliability when trained on severely underrepresented attack classes.

3. Real-Time Inference Performance Evaluation

The real-time inference performance of the proposed framework in terms of processing speed and efficiency. The model demonstrated a single-flow latency of 1.00 millisecond, enabling near-instantaneous classification of individual network flows. During batch evaluation, 100,000 flows were processed in just 0.94 seconds, resulting in a throughput of 106,630 flows per second and an average per-flow processing time of 0.0094 milliseconds. These results were obtained using a single-core CPU, highlighting the model's computational efficiency without relying on specialised hardware. The findings confirm the framework's capability to support real-time or near-real-time intrusion detection in practical deployment scenarios.

Table 3. Real-Time Inference Performance

Metric	Value
Single-flow latency	1.00 ms/flow
Batch inference time (100,000 flows)	0.94 sec
Throughput	106,630 flows/sec
Per-flow time (batch mode)	0.0094 ms/flow
Hardware	Single-core CPU

Source: Sharafaldin et al. (2018)

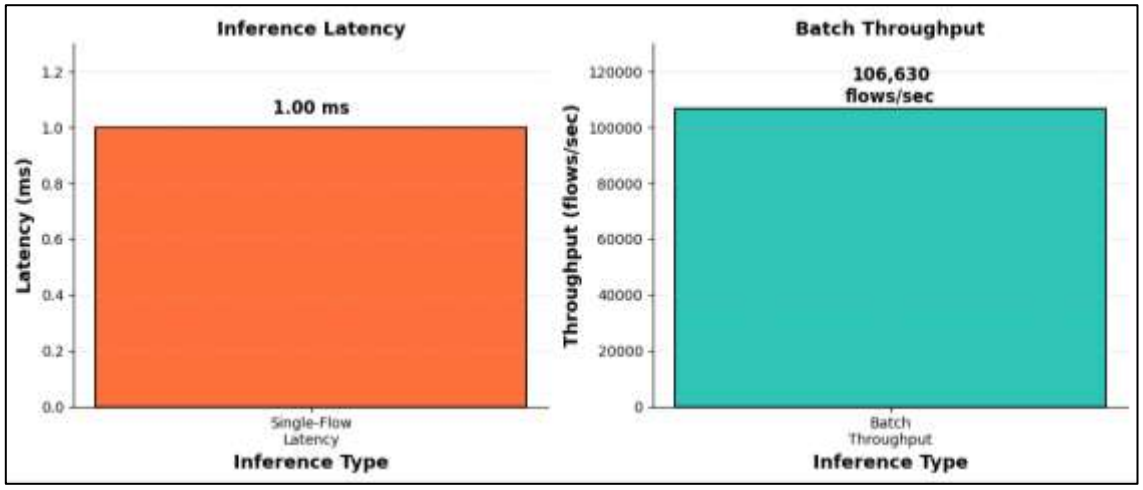

Figure 3. Real-Time Inference Performance (Single-Core CPU)

The real-time inference capability of the proposed framework using two performance dimensions: latency and throughput. The single-flow inference latency was measured at 1.00 milliseconds, indicating that the model can classify an individual network flow almost instantaneously, as shown in Figure 3. On the throughput side, the batch inference test processed traffic at a rate of 106,630 flows per second, demonstrating the model's capacity to handle high-volume network data efficiently. These results were obtained using a single-core CPU, suggesting that even greater performance could be achieved with additional computational resources. Together, these findings confirm the framework's suitability for real-time deployment scenarios.

Table 4. Framework Summary

Component	Result
Dataset	CICIDS2017 (2,830,743 flows, 78 features)
Anomaly Detection Accuracy	99.90%
Cyberattack Prediction Accuracy	98.64%
Real-time Throughput	106,630 flows/sec
Best Classes	DDoS, DoS Hulk, PortScan (F1 > 0.99)
Weakest Classes	Bot, Web Attacks, Infiltration

Source: Sharafaldin et al. (2018)

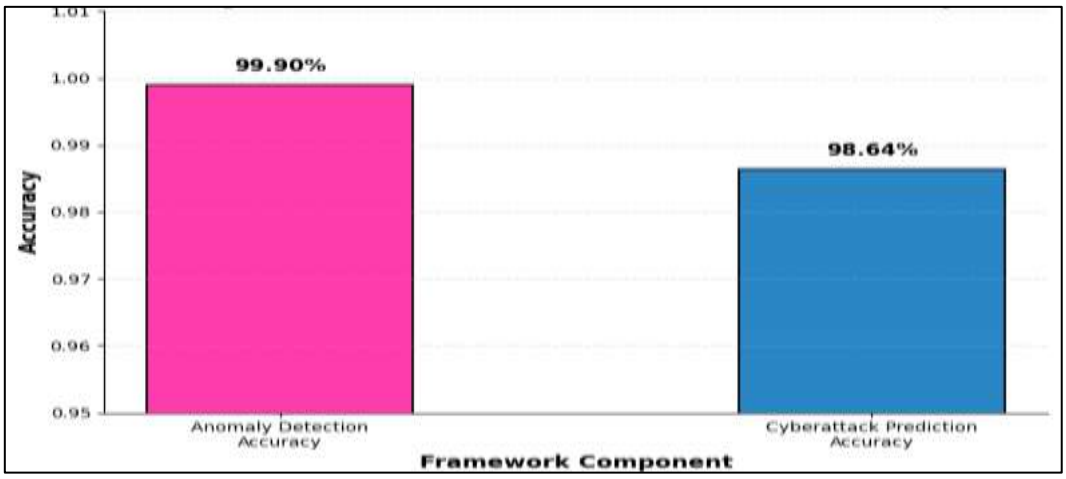

Figure 4. Overall Framework Performance Summary

The overall performance of the proposed framework by comparing its two core components. The anomaly detection module achieved an accuracy of 99.90%, reflecting its strong capability to distinguish benign traffic from anomalous behaviour, as shown in Figure 4. The cyberattack prediction module, responsible for identifying specific attack types, recorded a slightly lower accuracy of 98.64%, which is expected given the added complexity of multi-class classification. The marginal difference between the two accuracies

indicates that the framework maintains high reliability even when tasked with fine-grained attack categorisation. These results validate the effectiveness of the combined detection and prediction pipeline.

4. Discussion

The proposed machine-learning framework demonstrates strong overall performance across both core objectives of the study: anomaly detection and cyberattack-type classification. The binary anomaly detection model achieved an accuracy of 99.90%, with a precision of 0.9967 and recall of 0.9984 for the anomaly class, indicating that the framework distinguishes benign traffic from anomalous traffic with very few misclassifications. The correspondingly low false positive rate (0.082%) and false negative rate (0.162%) suggest the model would generate minimal unnecessary alerts while missing very few genuine threats a balance that matters in operational monitoring, where excessive false alarms lead to analyst fatigue and desensitisation to warnings.

The multi-class results present a more nuanced picture. High-frequency attack types such as DDoS, DoS Hulk, and PortScan were identified with near-perfect F1-scores (>0.99), but performance degraded substantially for rare categories, including Bot, Web Attack variants, Infiltration, and Heartbleed, where F1-scores fell to 0.17–0.24. This indicates that the framework's predictive capability is driven more by training-example volume per class than by the inherent detectability of the attack itself. The overall multi-class accuracy of 98.64% is therefore somewhat misleading in isolation, as it is dominated by majority classes and does not reflect the framework's weak sensitivity to rare but potentially high-impact attacks.

The real-time inference results, a single-flow latency of 1.00 ms and a throughput of 106,630 flows per second on a single-core CPU, indicate that computational overhead is low enough to support near-real-time traffic processing without specialised hardware. This directly supports the "real-time" component of the framework's objective, as the measured latency is well within the sub-second window generally required for practical intrusion-detection deployment.

The anomaly detection accuracy achieved here (99.90%) is consistent with, and in some respects exceeds, accuracy levels reported in related real-time anomaly detection research applied to critical infrastructure and IoT contexts. Sawas and Farag (2023) proposed a real-time anomaly prediction approach for detecting stealthy IoT-based cyber-attacks on power distribution systems, emphasising early and continuous detection in infrastructure-critical environments, a priority aligned with the low false-negative rate observed in the present framework. Similarly, Tayseer *et al*. (2025) developed a cyber-resilient machine-learning framework combining load forecasting with anomaly detection in smart grids, demonstrating that ML-based approaches achieve high detection reliability within operational monitoring pipelines, broadly consistent with the precision and recall values reported here.

The real-time findings also align with the direction of recent AI-driven cybersecurity research. Vignes *et al*. (2025) proposed an AI-driven cybersecurity framework for anomaly detection in power systems, highlighting the growing emphasis on intelligent detection under near-real-time constraints, consistent with the sub-millisecond latency achieved in this study. Zahid and Bharati (2025) similarly demonstrated that hybrid deep-learning approaches achieve effective real-time attack detection in IoT systems, reinforcing the trend toward combining detection accuracy with computational efficiency, an emphasis mirrored in this framework's dual evaluation of classification metrics and inference speed.

A notable divergence emerges, however, in the handling of class imbalance and rare-attack detection. Several cited studies focus on power-system-specific anomaly detection, where attack types are more narrowly defined, whereas this study evaluates a broader, more imbalanced set of network attack categories from a general-purpose intrusion detection dataset. The weaker performance observed for rare classes (F1-scores of 0.17–0.24) highlights a limitation not always explicitly addressed in prior work, suggesting generalised network-traffic frameworks face greater imbalance challenges than domain-specific systems trained on narrower attack taxonomies (VM *et al*., 2025).

These findings carry practical implications. First, the combination of high accuracy and low-latency inference suggests such frameworks could be integrated into operational monitoring without specialised, high-cost infrastructure, making them accessible to resource-constrained organisations. Second, the disparity between high-frequency and rare-attack performance implies that relying on aggregate accuracy alone could create a false sense of security regarding rare but severe attacks such as infiltration, underscoring the need to report per-class metrics rather than overall accuracy alone. Third, the demonstrated real-time capability supports the broader shift from reactive, offline detection toward continuously operating, low-latency monitoring pipelines suitable for dynamic network environments.

Several limitations should be acknowledged. The framework was evaluated using a single algorithm rather than a comparative set of models, limiting generalisability claims. Severe class imbalance, particularly for Heartbleed and Infiltration, was not corrected through resampling or class-weighting, likely contributing to low rare-class F1-scores. No formal feature selection or dimensionality reduction was applied, potentially introducing redundant variables into the learning process. Validation relied on a single

stratified train-test split rather than cross-validation, limiting robustness of the reported estimates. Finally, evaluation used a static, pre-collected dataset rather than live traffic, meaning the "real-time" claim reflects inference speed on historical data rather than validated performance under live, evolving deployment conditions.

Future work should compare multiple machine-learning and deep-learning algorithms to assess whether the observed performance generalises across different modelling approaches. Addressing class imbalance through oversampling, cost-sensitive learning, or ensemble-based rebalancing should be prioritised to improve detection of rare but critical attack types. Incorporating structured feature selection or engineering procedures may further improve model efficiency and interpretability. Future studies should also evaluate the framework under live or simulated streaming traffic conditions, rather than static datasets, to more rigorously validate real-time deployment claims and assess robustness against concept drift. Finally, extending the framework toward domain-specific applications such as IoT, industrial control systems, or smart-grid environments would help determine its generalisability beyond general-purpose network intrusion datasets and strengthen its applicability to critical infrastructure security.

5. Conclusion

This study developed a machine learning-based framework integrating real-time anomaly detection with cyberattack-type prediction, evaluated using a large-scale, publicly available network intrusion dataset containing multiple categories of benign and malicious traffic. The framework demonstrated strong capability in distinguishing benign traffic from anomalous behaviour, with consistently high precision and recall. Multi-class cyberattack prediction performed particularly well on high-frequency, well-represented attack types, though detection reliability declined noticeably for rare and underrepresented attack categories. Real-time inference testing confirmed that the framework could process network traffic with minimal latency and substantial throughput using standard computational resources. The study contributes an integrated analytical pipeline that connects anomaly detection with attack-type classification within a unified evaluation framework, rather than treating these as separate tasks. It further demonstrates that reliable detection performance and computational efficiency can be achieved without specialised or high-cost infrastructure, supporting the broader shift toward proactive, continuously operating cybersecurity systems. These findings suggest the framework holds practical value for deployment in resource-constrained network environments, offering organisations a computationally efficient approach to intrusion monitoring. The observed disparity in performance across attack categories underscores the importance of evaluating models using detailed, class-level metrics rather than relying solely on aggregate accuracy when assessing real-world deployment readiness. Future research should incorporate comparative modelling across multiple algorithms, address underlying class imbalance through appropriate technical strategies, and validate framework performance under live streaming traffic conditions and domain-specific environments such as IoT and industrial control systems, thereby strengthening the generalisability and practical applicability of the proposed approach.

References

1. Abokifa, A. A., Haddad, K., Lo, C., & Biswas, P. (2019). Real-time identification of cyber-physical attacks on water distribution systems via machine learning–based anomaly detection techniques. *Journal of Water Resources Planning and Management*, *145*(1), 04018089.
2. Alabdulatif, A. A., Thilakarathne, N. N., & Aashiq, M. (2024). Machine Learning-Enabled Novel Real-Time IoT Targeted DoS/DDoS Cyber Attack Detection System. *Comput. Mater. Continua*, *80*(3), 3655-3683.
3. Almalki, S. S. (2025). A Deep Learning-Based Framework for Real-Time Detection of Cybersecurity Threats in IoT Environments. *International Journal of Advanced Computer Science & Applications*, *16*(3), 430.
4. Alsbatin, L., Zawaideh, F., Alrifai, B. M., & Alawneh, T. A. (2025). Enhancing Internet of Things (IoT) Network Security: A Machine Learning-Driven Framework for Real-Time Intrusion Detection and Anomaly Classification. *Mesopotamian Journal of CyberSecurity*, *5*(3), 1042-1056.
5. Elnour, M., Noorizadeh, M., Shakerpour, M., Meskin, N., Khan, K., & Jain, R. (2023). A machine learning based framework for real-time detection and mitigation of sensor false data injection cyber-physical attacks in industrial control systems. *IEEe Access*, *11*, 86977-86998.
6. Gulzar, Q., & Mustafa, K. (2025). Interdisciplinary framework for cyber-attacks and anomaly detection in industrial control systems using deep learning. *Scientific Reports*, *15*(1), 26575.
7. Hao, W., Yang, T., & Yang, Q. (2021). Hybrid statistical-machine learning for real-time anomaly detection in industrial cyber–physical systems. *IEEE Transactions on Automation Science and Engineering*, *20*(1), 32-46.

8. Jain, V., & Mitra, A. (2025). Real-time threat detection in cybersecurity: leveraging machine learning algorithms for enhanced anomaly detection. In *Machine Intelligence Applications in Cyber-Risk Management* (pp. 315-344). IGI Global Scientific Publishing.
9. Kabir, S., Hannan, N., Shufian, A., & Zishan, M. S. R. (2025). Proactive detection of cyber-physical grid attacks: A pre-attack phase identification and analysis using anomaly-based machine learning models. *Array*, *27*, 100441.
10. Kavousi-Fard, A., Su, W., & Jin, T. (2020). A machine-learning-based cyber attack detection model for wireless sensor networks in microgrids. *IEEE Transactions on Industrial Informatics*, *17*(1), 650-658.
11. Kesici, M., Mohammadpourfard, M., Aygul, K., & Genc, I. (2023). Deep learning-based framework for real-time transient stability prediction under stealthy data integrity attacks. *Electric Power Systems Research*, *221*, 109424.
12. Komisarek, M., Pawlicki, M., Kozik, R., & Choras, M. (2021). Machine Learning Based Approach to Anomaly and Cyberattack Detection in Streamed Network Traffic Data. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, *12*(1), 3-19.
13. Mizanur, M., Kumer, S., & Reza, N. (2025). Machine learning-based anomaly detection for cyber threat prevention. *Journal of Primeasia*, *6*(1), 1-8.
14. Naveeda, K., & Fathima, S. S. S. (2025). Real-time implementation of IoT-enabled cyberattack detection system in advanced metering infrastructure using machine learning technique. *Electrical Engineering*, *107*(1), 909-928.
15. Raje, V. V., Goel, S., Patil, S. V., Kokate, M. D., Mane, D. A., & Lavate, S. (2023). Realtime Anomaly Detection in Healthcare IoT: A Machine Learning-Driven Security Framework. *Journal of Electrical Systems*, *19*(3), 192.
16. Raza, A., Ali, A. K. S., & Hussain, A. A. (2024). AI-driven approaches to cyber and information security: Machine learning algorithms for threat prediction and anomaly detection. *Spectrum of Engineering Sciences*, 565-573.
17. Saeed, M. M. (2025). An AI-driven cybersecurity framework for IoT: integrating LSTM-based anomaly detection, reinforcement learning, and post-quantum encryption. *IEEE Access*, *13*, 104027-104036.
18. Salman, A. M., Al-Nuaimi, B. T., Subhi, A. A., & Alkattan, H. (2025). Enhancing cybersecurity with machine learning: A hybrid approach for anomaly detection and threat prediction. Mesopotamian Journal of CyberSecurity, 5(1), 202-215.
19. Sawas, A., & Farag, H. E. (2023). Real-time detection of stealthy IoT-based cyber-attacks on power distribution systems: A novel anomaly prediction approach. *Electric Power Systems Research*, *223*, 109496.
20. Tayseer, M., Talaat, M., Zamel, A. A., Sedhom, B. E., Elgamal, M., Senjyu, T., ... & Elkholy, M. H. (2025). Cyber-resilient machine learning framework for accurate individual load forecasting and anomaly detection in smart grids. *Scientific Reports*, *15*(1), 44054.
21. Vignes, V. M., MP, S. H., Satheesh, R., Das, V., & Padmanaban, S. (2025). AI-driven cybersecurity framework for anomaly detection in power systems. *Scientific Reports*, *15*, 35506.
22. VM, V., MP, S. H., Satheesh, R., Das, V., & Padmanaban, S. (2025). Ai-driven cybersecurity framework for anomaly detection in power systems: Vignes vm et al. *Scientific Reports*, *15*(1), 35506.
23. Zahid, M., & Bharati, T. S. (2025). Enhancing cybersecurity in IoT systems: a hybrid deep learning approach for real-time attack detection. *Discover Internet of Things*, *5*(1), 73.
24. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, *1*(2018), 108-116.
25. Ahmed, C. M., MR, G. R., & Mathur, A. P. (2020, October). Challenges in machine learning based approaches for real-time anomaly detection in industrial control systems. In *Proceedings of the 6th ACM on cyber-physical system security workshop* (pp. 23-29).
26. Goel, N. (2025, October). Enhancing Cybersecurity Frameworks Using Artificial Intelligence and Deep Learning for Real-Time Threat Detection and Prevention. In *2025 2nd International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)* (pp. 1-7). IEEE.