World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

DESIGN AND IMPLEMENTATION OF AN INTELLIGENT HYBRID INTRUSION DETECTION SYSTEM USING SIGNATURE-BASED DETECTION, LIGHTGBM, AND ISOLATION FOREST

Kingsley I. Chibueze [1,*], C.O. Ugwoke [2], Emeka Augustine Chinachi [3], Malachy O. Ugwuoke [4], Onyeabo Uzoamaka Agatha [5] and Okika Stephen Sunday [6]

[1] Department of Computer Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
[2] Department of Civil Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
[3] Department of Mechanical Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
[4] Department of Chemical Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
[5] Department of Biomedical Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
[6] Department of Electrical and Electronic Engineering, Faculty of Engineering and the Built Environment, State University of Medical and Applied Sciences, Igbo-Eno, Enugu State, Nigeria.
* Corresponding Author

ORCID Details
Kingsley I. Chibueze: https://orcid.org/0009-0008-9387-0297

World Journal of Advanced Research and Reviews, 2026, 31(03), 980–988

Publication history: Received on 02 August 2026; revised on 13 September 2026; accepted on 15 September 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.31.3.2404

ABSTRACT

Network intrusion detection systems must identify known attacks while remaining sensitive to previously unseen or behaviorally unusual traffic. Single detection paradigms are limited because signature-based methods depend on known attack patterns, supervised classifiers may miss behavior outside their training distribution, and anomaly detectors may generate false alarms when legitimate traffic changes. This study presents an intelligent hybrid network intrusion detection system that combines signature-based detection, LightGBM supervised classification, and Isolation Forest anomaly detection in a sequential pipeline. The system was implemented as a Python/Flask application with passive packet monitoring and manual analysis of JSON/CSV traffic records. HIKARI-2021 was used for supervised training. The dataset contained 555,278 records; after preprocessing, 81 features were retained. A stratified 80:20 train-test split was used for LightGBM, while Isolation Forest was trained on 414,065 benign training records. LightGBM was configured with 250 estimators, a learning rate of 0.05, 48 leaves, class weighting, and a tuned decision threshold of 0.7987. On the test set, the classifier achieved 90.60% accuracy, 94.63% weighted precision, 90.60% weighted recall, and 92.01% weighted F1-score. The suspicious-class recall was 81.89%, while suspicious-class precision was 40.50%. The ROC curve produced an AUC of 0.9522. The results demonstrate that the supervised component provides useful

discrimination, while the hybrid architecture supplies complementary signature and anomaly-detection mechanisms. Because the available experiment did not produce a separate end-to-end quantitative evaluation of the complete hybrid pipeline, the reported numerical metrics are explicitly attributed to the LightGBM component.

Keywords: Hybrid Intrusion Detection; Network Intrusion Detection; Lightgbm; Isolation Forest; Signature-Based Detection; HIKARI-2021; SHAP; Explainable Artificial Intelligence.

1 INTRODUCTION

Network intrusion detection systems (IDSs) are an important layer of network security because they monitor traffic and identify activities that may compromise confidentiality, integrity, or availability. Two established detection paradigms are misuse/signature-based detection and anomaly-based detection. Signature-based detection is effective when the characteristics of an attack are already represented by known rules, but it is inherently constrained when an attack is modified or previously unseen. Anomaly-based detection addresses a different part of the problem by learning characteristics of expected behavior and flagging deviations. However, deviations caused by legitimate changes in traffic can also produce false alarms.

Hybrid IDS architectures attempt to combine complementary detection mechanisms. Maseno et al. (2022) reviewed 111 studies of hybrid IDS research and identified the combination of multiple detection approaches as a means of exploiting the strengths of individual methods while reducing their weaknesses. Their review also emphasizes that detection performance alone is not sufficient for practical deployment; real-world systems must address evolving traffic, false positives, computational requirements, and operational usability.

Machine learning has become central to network-flow intrusion detection because it can learn nonlinear relationships among traffic features. LightGBM is particularly suitable for structured tabular data and has been applied to intrusion detection in several recent studies. For example, Chen et al. (2023) used LightGBM as the first stage of a two-stage IDS, while Dalal et al. (2023) optimized LightGBM using Bayesian optimization for cyber-physical-system intrusion detection. These studies demonstrate the capability of LightGBM, but they do not combine it with the signature-plus-unsupervised-anomaly sequence developed in the present work.

The choice of HIKARI-2021 is also significant. Ferriyan et al. (2021) introduced HIKARI-2021 to provide a contemporary intrusion-detection benchmark containing benign traffic and encrypted synthetic attack traffic. The dataset is therefore more aligned with modern encrypted network environments than many older IDS benchmarks. Nevertheless, it was generated in a controlled laboratory setting, so results obtained from it should not be interpreted as a complete substitute for evaluation on live operational networks.

A further practical issue is the interpretability of machine-learning alerts. Security analysts need more than a binary prediction when deciding whether an alert represents a genuine incident. Recent explainable-IDS research has emphasized the importance of producing understandable explanations for model decisions (Neupane et al., 2022), while Gaspar et al. (2024) demonstrated the usefulness of SHAP and LIME for interpreting IDS predictions.

This study therefore focuses on a single integrated contribution: a practical hybrid IDS in which deterministic signature inspection is followed by LightGBM classification and, for records classified as benign, an Isolation Forest behavioral check. SHAP is used to support interpretation of suspicious machine-learning decisions. The principal contribution is the integration of these mechanisms into one deployable workflow rather than a claim that any individual classifier is state-of-the-art.

The objectives were to: (i) design a sequential hybrid detection architecture; (ii) implement the architecture as a working web-based IDS; (iii) train and evaluate the LightGBM component on HIKARI-2021; and (iv) integrate an unsupervised anomaly layer and explanation mechanism into the operational workflow.

2 LITERATURE REVIEW

Signature-based detection identifies malicious traffic by matching observable characteristics against predefined rules. Its principal advantage is efficient identification of known attacks with relatively direct decision logic. Its principal limitation is that a new or modified attack may not match an existing rule. Hybrid IDS research has consequently explored combining signature or misuse detection with anomaly-oriented techniques.

Anomaly detection instead models normal or expected behavior and identifies observations that deviate from it. Isolation Forest is an unsupervised approach that isolates observations through randomized partitioning; anomalous observations tend to be isolated more quickly than observations belonging to dense normal regions. Al-Shehari et al.

(2023) demonstrated the use of Isolation Forest for insider-threat detection, supporting its use as a complementary behavioral detector rather than as a replacement for supervised classification.

Maseno et al. (2022) found that hybrid IDS research commonly uses cascaded or integrated combinations of detection algorithms. A cascaded design is particularly relevant to the present study because it permits different mechanisms to process the same traffic at different decision stages. The proposed system adopts this principle by placing the low-cost signature check first, supervised classification second, and unsupervised anomaly analysis third.

Recent hybrid machine-learning studies have used different forms of model combination. Chen et al. (2023) combined LightGBM and a convolutional neural network in a two-stage architecture, while Ayubkhan et al. (2023) combined a denoising autoencoder with LightGBM to improve robustness to noisy network traffic. These studies show the value of complementary stages, but neither implements the specific signature-first, LightGBM, Isolation Forest sequence evaluated here.

LightGBM is a gradient-boosting framework designed for efficient learning on structured data. Its use in IDS is attractive because network-flow datasets contain numerous numerical and categorical features and can be large enough to make computational efficiency important. Dalal et al. (2023) reported 99.17% accuracy on UNR-IDD after Bayesian optimization, while Chen et al. (2023) reported that LightGBM trained faster than several comparison classifiers in their two-stage IDS. These results motivate the selection of LightGBM as the supervised component of the present hybrid architecture, but the performance values are not directly comparable because the datasets, preprocessing procedures, class distributions, and evaluation protocols differ.

Ferriyan et al. (2021) introduced HIKARI-2021 as an IDS dataset containing benign traffic and encrypted synthetic attacks, with procedures for reproducing and extending the dataset. Ribeiro et al. (2022) subsequently examined machine-learning intrusion detection using HIKARI-2021 and reported that feature selection could substantially reduce the feature space while retaining strong classification performance.

Explainability is increasingly treated as an operational requirement for IDS rather than merely a visualization feature. Neupane et al. (2022) surveyed explainable IDS research and identified the need for explanations that are meaningful to different stakeholders. Gaspar et al. (2024) compared SHAP and LIME in an IDS context and used perturbation analysis to examine explanation behavior. Chen et al. (2024) similarly used SHAP-guided feature selection for encrypted-traffic intrusion detection, although their evaluation focused on CICIDS2017 and NSL-KDD rather than HIKARI-2021.

2.1 Research Gap

The reviewed literature demonstrates substantial progress in hybrid detection, LightGBM-based classification, anomaly detection, and explainable IDS. However, these capabilities are usually investigated separately or in different combinations. The gap addressed by this study is therefore architectural and operational: the implementation of a single sequential IDS that begins with signature inspection, applies LightGBM to unresolved traffic, subjects LightGBM-benign traffic to an Isolation Forest check, and provides SHAP-supported interpretation for suspicious machine-learning outputs. The contribution is consequently the integration and implementation of these mechanisms in one workflow, together with an evaluation of the supervised LightGBM component on HIKARI-2021.

3 MATERIALS AND METHODS

3.1 Research and System Development Approach

The study followed an experimental and systems-development approach. Object-Oriented Analysis and Design (OOAD) was used to structure the application into modular services for data processing, signature checking, supervised classification, anomaly detection, real-time monitoring, logging, and reporting. The implementation was developed in Python with Flask. Scapy supported passive packet capture for real-time monitoring, while SHAP was integrated for model interpretation.

The system accepts manually supplied JSON/CSV traffic records and can observe traffic passively through a network interface. The detection sequence was designed so that deterministic checks are performed before computationally heavier model-based analysis.

3.2 Hybrid Detection Architecture

The architecture follows a sequential pipeline. First, an incoming traffic record is normalized and inspected by the signature engine. A matching signature produces an immediate signature alert. If no signature is matched, the record is

mapped to the HIKARI-2021 feature schema and passed to the LightGBM classifier. If LightGBM predicts suspicious traffic, the system generates an explanation using SHAP where available. If LightGBM predicts benign traffic, the record is passed to Isolation Forest for a second behavioral check. This arrangement is intended to capture known attacks through deterministic rules, learned malicious patterns through supervised classification, and unusual benign-looking behavior through unsupervised analysis.

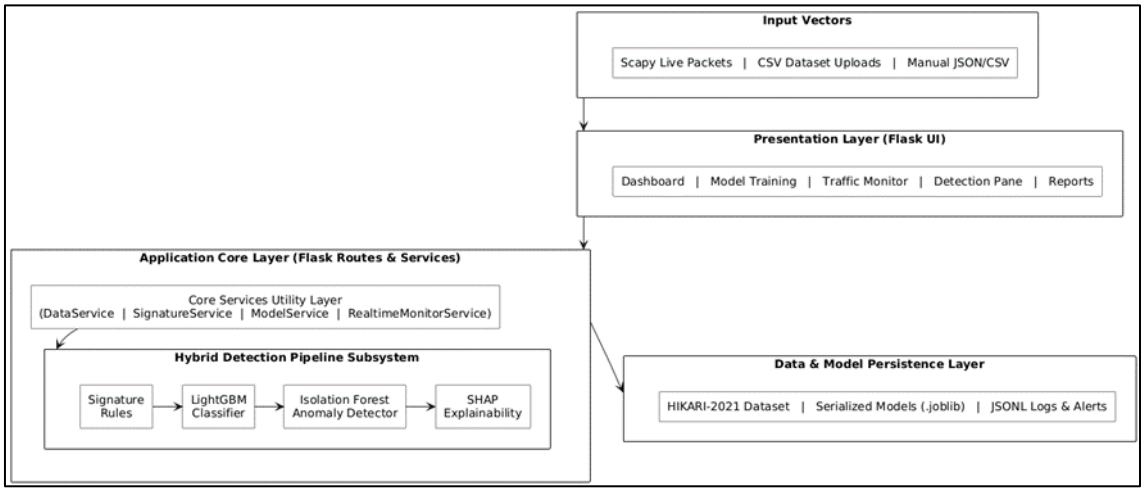

Figure 1: System architecture of the proposed intelligent hybrid intrusion detection system.

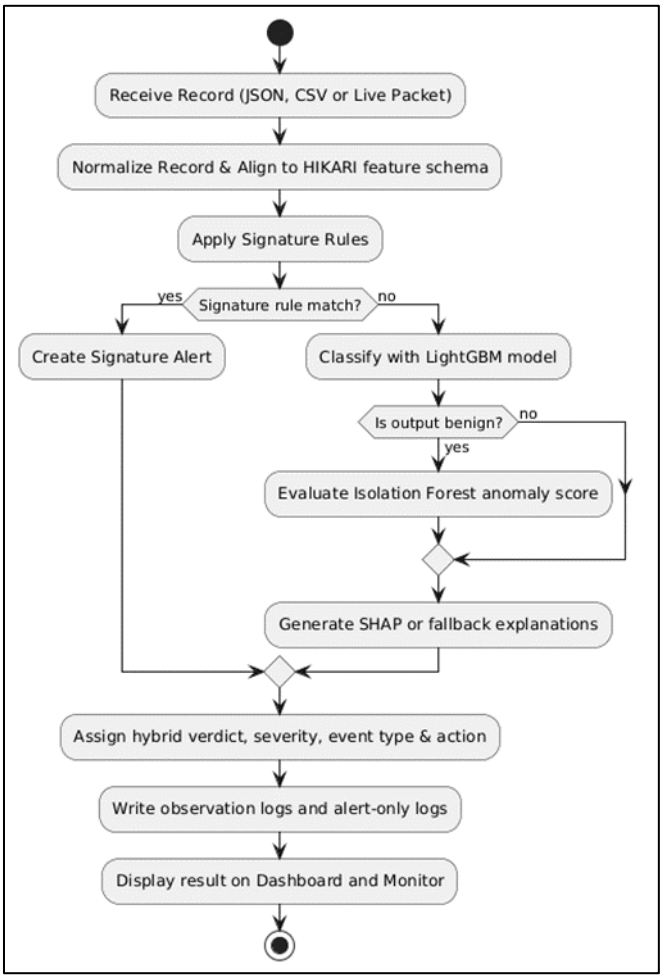

Figure 2: Activity workflow of the sequential hybrid detection process.

The operational sequence is illustrated in Figure 2. The workflow terminates in a logged detection result containing the decision and supporting information for analyst review.

3.3 Dataset and Preprocessing

HIKARI-2021 was used as the principal training dataset. The source project reports 555,278 records and an original feature space of 83 columns. The binary target variable was Label, with 0 representing normal traffic and 1 representing suspected traffic. For the implemented model, identifier-like and context-specific fields were removed, including Unnamed: 0, Unnamed: 0.1, originh, responh, traffic_category, and uid, leaving 81 usable features.

Missing numerical values were median-imputed. Missing categorical values were replaced by the mode and ordinal-encoded. The supervised model used an 80:20 stratified train-test split, giving 444,222 training records and 111,056 test records. Isolation Forest was trained using 414,065 benign training records. Figure 3 summarizes these quantities.

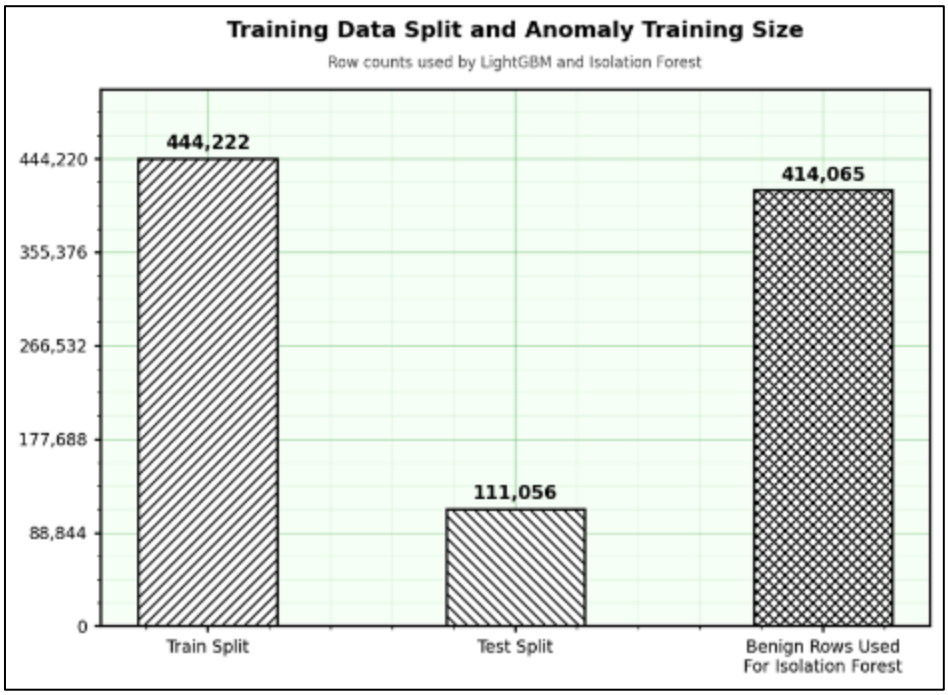

Figure 3: Training-test split and benign records used to train the Isolation Forest component.

3.4 LightGBM Model Configuration

The LightGBM classifier used 250 estimators, a learning rate of 0.05, and 48 leaves. Class weighting was applied to reduce the effect of class imbalance. The decision threshold was tuned to 0.7987. The supervised model was evaluated using accuracy, weighted precision, weighted recall, weighted F1-score, class-specific precision, class-specific recall, class-specific F1-score, confusion matrix values, and the ROC curve.

3.5 Isolation Forest Configuration

Isolation Forest was trained only on benign traffic so that its learned representation corresponded to normal behavior. The implementation used 200 estimators, a contamination value of 0.05, and an anomaly alert threshold of 0.51.

3.6 Explainability and Detection Output

SHAP was integrated to provide feature-level explanations for suspicious machine-learning decisions. Where SHAP output was unavailable, the application could produce a fallback textual explanation based on traffic characteristics such as protocol, port numbers, duration, TCP flag counts, packet rate, and payload volume. This explanation mechanism is treated as an operational aid rather than as a separately evaluated XAI contribution; quantitative explanation-quality evaluation is reserved for future work.

4 RESULTS AND DISCUSSION

4.1 LightGBM Classification Performance

The LightGBM classifier achieved 90.60% accuracy on the HIKARI-2021 test split. Weighted precision was 94.63%, weighted recall was 90.60%, and weighted F1-score was 92.01%. For the suspicious class, precision was 40.50%, recall

was 81.89%, and F1-score was 54.20%. The relatively high suspicious-class recall is important for intrusion detection because missed attacks can be more consequential than additional investigation caused by false alarms.

The ROC curve produced an AUC of 0.9522, indicating strong separation between the two classes across decision thresholds. However, the AUC should not be interpreted in isolation: the suspicious-class precision of 40.50% shows that a substantial number of benign records were still flagged as suspicious at the selected operating point.

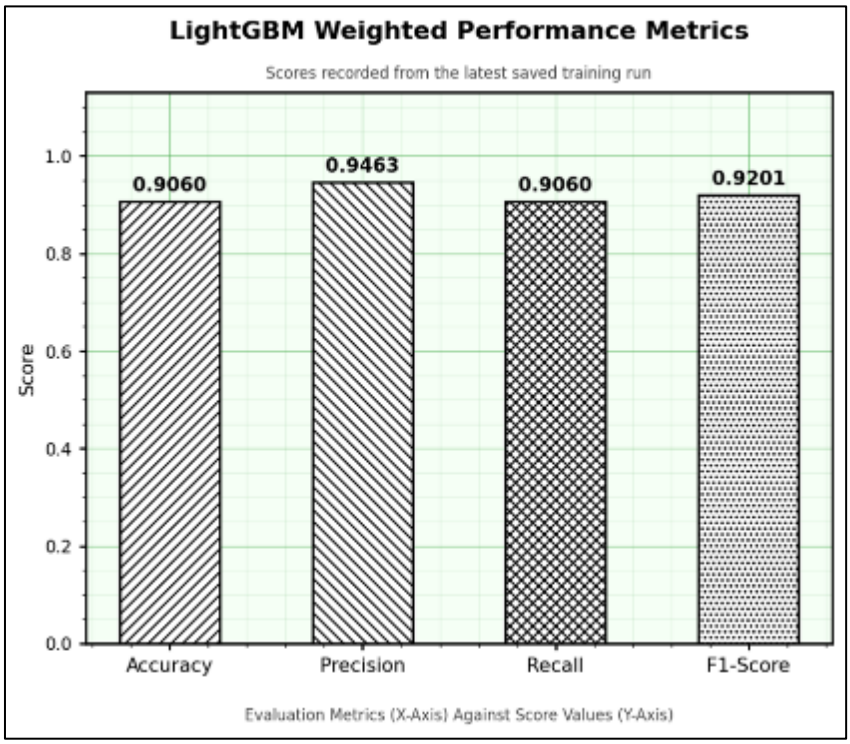

Figure 4: Weighted accuracy, precision, recall, and F1-score obtained by the LightGBM classifier.

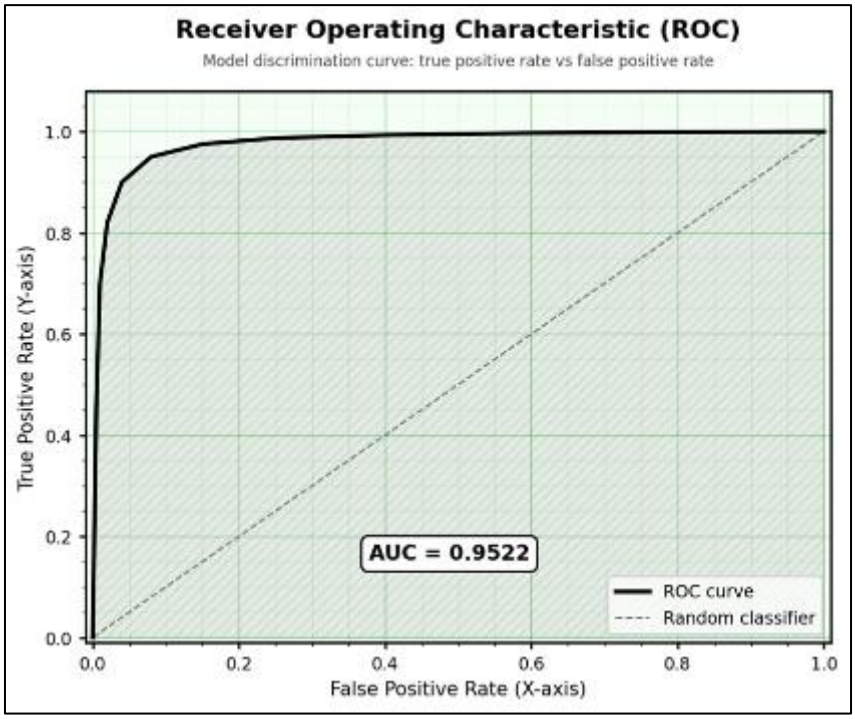

Figure 5: Receiver operating characteristic curve for the LightGBM classifier (AUC = 0.9522).

4.2 Confusion-Matrix Analysis

The confusion matrix provides a more detailed view of the classifier behavior. Of the benign test records, 94,446 were correctly classified as benign and 9,071 were incorrectly classified as suspicious. Among suspicious records, 6,174 were correctly detected and 1,365 were classified as benign. The resulting false-negative rate for the suspicious class is therefore lower than the corresponding false-positive burden, which is consistent with the reported suspicious-class recall of 81.89% and precision of 40.50%.

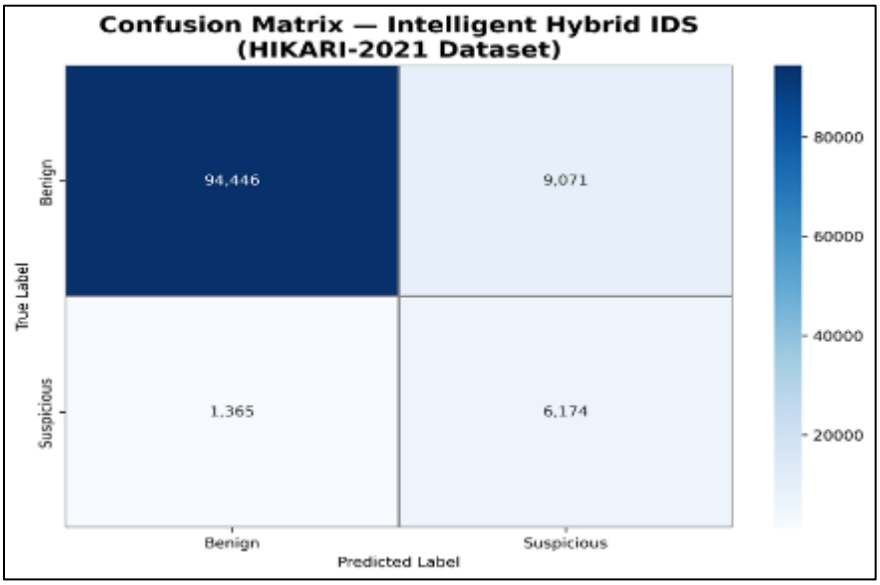

Figure 6: Confusion matrix for the LightGBM classifier on the HIKARI-2021 test set.

4.3 Interpretation of the Hybrid System

The most important system-level result is architectural rather than a single accuracy value. The implemented system places signature detection ahead of machine learning, allowing known patterns to be identified without requiring a model inference step. Traffic that does not match a signature is then evaluated by LightGBM. Records classified as benign are subjected to an additional Isolation Forest check, creating a route through which behaviorally unusual records can still be escalated. Suspicious machine-learning decisions can be accompanied by SHAP-supported explanations, improving the information available to an analyst.

This design is consistent with the hybrid-IDS literature, in which cascaded approaches are used to combine complementary detection mechanisms (Maseno et al., 2022). It also addresses a limitation identified in studies that focus only on supervised classification. For example, Chen et al. (2023) demonstrated strong performance with a LightGBM-plus-CNN architecture, but the present system addresses a different objective by adding deterministic signatures, an unsupervised behavioral stage, and analyst-oriented explanation.

The quantitative results must nevertheless be interpreted carefully. The reported 90.60% accuracy, 94.63% weighted precision, 90.60% weighted recall, 92.01% weighted F1-score, and 0.9522 AUC are measurements of the evaluated LightGBM classifier. The original experiment did not provide a separately labeled end-to-end evaluation in which the signature, LightGBM, and Isolation Forest stages were jointly scored. Therefore, these values should not be presented as the accuracy of the complete hybrid system.

4.4 Comparison with Recent Studies

Direct numerical comparison across IDS studies should be treated cautiously because datasets, class distributions, preprocessing, thresholds, and evaluation protocols differ. The more useful comparison is therefore the combination of capabilities addressed by each study.

Table 1: Comparison with Recent Studies

Study	Main approach	Principal result/contribution	Key distinction or limitation
Maseno et al. (2022)	Systematic review of hybrid IDS	111 hybrid-IDS studies reviewed; hybrid approaches address weaknesses of individual methods.	Review study; no implementation of the present architecture.
Chen et al. (2023)	LightGBM + CNN, CSE-CIC-IDS2018	High-performance two-stage supervised IDS.	No signature-first + Isolation Forest + SHAP workflow.
Ayubkhan et al. (2023)	Denoising autoencoder + LightGBM	Improved robustness to noisy traffic.	Different hybridization objective; no signature/anomaly/XAI pipeline.
Al-Shehari et al. (2023)	Isolation Forest	Demonstrated anomaly-based insider-threat detection.	Does not provide the supervised/signature combination used here.
Dalal et al. (2023)	Bayesian-optimized LightGBM, UNR-IDD	99.17% accuracy reported on its evaluation dataset.	Single supervised model; different dataset and architecture.
Chen et al. (2024)	SHAP-guided feature selection	Interpretable feature selection for IDS.	Evaluated on CICIDS2017 and NSL-KDD; not the present hybrid pipeline.
Gaspar et al. (2024)	SHAP and LIME for IDS	Perturbation analysis of explanation behavior.	XAI-focused study rather than a multi-stage hybrid detector.
Present study	Signature → LightGBM → Isolation Forest + SHAP-supported output	90.60% accuracy; 94.63% weighted precision; 90.60% weighted recall; 92.01% weighted F1; AUC 0.9522 for LightGBM.	End-to-end hybrid performance requires separate quantitative evaluation.

Limitations

Three limitations should be emphasized. First, HIKARI-2021 is laboratory-generated and may not fully represent traffic distributions encountered in operational networks. Second, the suspicious-class precision of 40.50% indicates a non-trivial false-positive burden at the selected operating point. Third, the complete hybrid pipeline was not evaluated with a separately labeled end-to-end test set; consequently, the contribution of the Isolation Forest and signature stages cannot be summarized by a single experimentally verified hybrid accuracy from the available results.

5 CONCLUSION

This study designed and implemented an intelligent hybrid network intrusion detection system that integrates signature-based detection, LightGBM supervised classification, and Isolation Forest anomaly detection in a sequential workflow. The system was implemented as a Flask-based application with passive monitoring, manual traffic analysis, logging, analyst review, and SHAP-supported interpretation of suspicious machine-learning decisions.

On HIKARI-2021, the evaluated LightGBM component achieved 90.60% accuracy, 94.63% weighted precision, 90.60% weighted recall, 92.01% weighted F1-score, and an AUC of 0.9522. The suspicious-class recall of 81.89% indicates useful sensitivity to suspicious traffic, while the 40.50% suspicious-class precision shows that false-positive reduction remains an important challenge. The principal contribution of the work is therefore the integrated architecture and operational workflow, not a claim of superior standalone classification accuracy.

Suggestions for Future Work

- Evaluate the complete hybrid pipeline on additional contemporary network-flow datasets and on controlled real-network traffic.
- Perform a labeled end-to-end experiment that separately measures the contribution of the signature, LightGBM, and Isolation Forest stages.
- Evaluate Isolation Forest using anomaly-score distributions, threshold sensitivity, detection rate, false-alarm rate, and processing latency.

- Quantitatively evaluate SHAP explanations for fidelity, stability, computational overhead, and usefulness to security analysts.
- Investigate calibrated decision thresholds and class-imbalance strategies to reduce the false-positive burden while preserving suspicious-class recall.
- Measure real-time throughput, inference latency, memory use, CPU utilization, and SHAP explanation time under realistic traffic loads.
- Investigate adaptive signature updates and online or incremental learning so that the system can respond to evolving attack patterns.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

I wish to acknowledges all Co- authors for their contributions in this work.

Disclosure of conflict of interest

I declared that there is no conflict of interests exist.

REFERENCES

[1] Al-Shehari, T., Al-Razgan, M. S., Alfakih, T., Alsowail, R. A., & Pandiaraj, S. (2023). Insider threat detection model using anomaly-based Isolation Forest algorithm. IEEE Access, 11, 118170–118185. https://doi.org/10.1109/ACCESS.2023.3326750

[2] Ayubkhan, S. A. H., Yap, W.-S., Morris, E., & Rawthar, M. B. K. (2023). A practical intrusion detection system based on denoising autoencoder and LightGBM classifier with improved detection performance. Journal of Ambient Intelligence and Humanized Computing, 14(6), 7427–7452. https://doi.org/10.1007/s12652-022-04449-w

[3] Chen, X., Liu, M., Wang, Z., & Wang, Y. (2024). Explainable deep learning-based feature selection and intrusion detection method on the Internet of Things. Sensors, 24(16), 5223. https://doi.org/10.3390/s24165223

[4] Chen, Y., Li, Z., & Wang, H. (2023). A two-stage intrusion detection system using LightGBM and convolutional neural network. Expert Systems with Applications, 213, 119918. https://doi.org/10.1016/j.eswa.2023.119918

[5] Dalal, S., Poongodi, M., Lilhore, U. K., Dahan, F., Vaiyapuri, T., Keshta, I., Aldossary, S. M., Mahmoud, A., & Simaiya, S. (2023). Optimized LightGBM model for security and privacy issues in cyber-physical systems. Transactions on Emerging Telecommunications Technologies, 34(5), e4771. https://doi.org/10.1002/ett.4771

[6] Ferriyan, A., Thamrin, A. H., Takeda, K., & Murai, J. (2021). Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic. Applied Sciences, 11(17), 7868. https://doi.org/10.3390/app11177868

[7] Gaspar, D., Silva, P., & Silva, C. (2024). Explainable AI for intrusion detection systems: LIME and SHAP applicability on multi-layer perceptron. IEEE Access, 12, 30164–30175. https://doi.org/10.1109/ACCESS.2024.3368377

[8] Maseno, E. M., Wang, Z., & Xing, H. (2022). A systematic review on hybrid intrusion detection system. Security and Communication Networks, 2022, Article 9663052. https://doi.org/10.1155/2022/9663052

[9] Neupane, S., Ables, J., Anderson, W., Mittal, S., Rahimi, S., Banicescu, I., & Seale, M. (2022). Explainable intrusion detection systems (X-IDS): A survey of current methods, challenges, and opportunities. IEEE Access, 10, 112392–112415. https://doi.org/10.1109/ACCESS.2022.3216617

[10] Ribeiro, M., Protasio, I., Marques, A., & Gama, J. (2022). Machine learning-based intrusion detection using the HIKARI-2021 dataset. Proceedings of the IEEE International Conference on Cyber Security and Resilience. https://doi.org/10.1109/CSR54599.2022.9800807

[11] Wang, Z., Zheng, L., Yang, M., & Wang, M. (2020). Explainability of machine learning-based intrusion detection systems using SHAP. IEEE Access, 8, 126390–126401. https://doi.org/10.1109/ACCESS.2020.9069273