IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

AI-Based Dynamic Risk Scoring Improves Zero Trust Access Control Accuracy and Reduces False Positives

Abhishek Pal[1] and Dr. Ritu Pal[2]
[1]M.Tech Scholar, Department of Computer Science, Tula's Institute, Dehradun, Uttarakhand, India
[2]Department of Computer Science, Tula's Institute, Dehradun, Uttarakhand, India
[1]abhisinghpal3@gmail.com and [2]ritu.pal@tulas.edu.in

Abstract: *The Zero Trust security model abandons the notion of an implicitly trusted internal network and instead requires every access request to be continuously verified. Conventional Zero Trust deployments, however, rely heavily on static, rule-based policy engines that evaluate requests against fixed conditions. Such rigid rules generalise poorly to evolving threats, produce a large number of false alarms, and impose unnecessary friction on legitimate users. This paper proposes an Artificial Intelligence (AI) based dynamic risk-scoring framework that augments Zero Trust access control with a continuously updated, context-aware risk score. The framework extracts contextual features spanning user identity, device posture, network behaviour, geolocation and temporal patterns, and feeds them to a heterogeneous machine-learning ensemble combining Random Forest, Extreme Gradient Boosting (XGBoost) and a Long Short-Term Memory (LSTM) network. The ensemble outputs a normalised risk score R(t) in the range 0–100, which is compared against adaptive thresholds at the Policy Decision Point to grant, step-up (multi-factor) or deny access. The approach was evaluated on three widely used benchmark datasets—NSL-KDD, CICIDS2017 and UNSW-NB15—and benchmarked against a representative static rule-based baseline. Experimental results show that the proposed model raises detection accuracy to an average of 97.2% and an Area Under the Curve (AUC) of 0.984, while reducing the false positive rate from an average of 17.0% for the static baseline to 3.1%—a relative reduction of more than 80%. These results demonstrate that AI-driven dynamic risk scoring materially improves the accuracy of Zero Trust access decisions while substantially reducing false positives, thereby lowering operational burden without compromising security.*

Keywords: Zero Trust Architecture; Dynamic Risk Scoring; Machine Learning; Access Control; Adaptive Authentication; Intrusion Detection; False Positive Reduction; Cybersecurity; Ensemble Learning; Continuous Authentication

I. INTRODUCTION

Modern enterprises operate in an environment where the traditional security perimeter has effectively dissolved. The widespread adoption of cloud computing, mobile workforces, bring-your-own-device practices and distributed micro-service architectures means that users, devices and workloads routinely connect from outside the corporate boundary. In this setting, the classical castle-and-moat model—which grants broad trust to anything located inside the network—has become a serious liability. Once an attacker breaches the perimeter, lateral movement is often trivial because internal traffic is implicitly trusted.

The Zero Trust security model emerged as a direct response to these shortcomings. Its central tenet, frequently summarised as "never trust, always verify," mandates that no entity be trusted by default, regardless of whether it sits inside or outside the network. Every access request must be authenticated, authorised and continuously validated

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, June 2026 Impact Factor: 8.2

against the security posture of the requesting subject before access to a resource is granted. The National Institute of Standards and Technology (NIST) formalised these principles in Special Publication 800-207, which describes a logical architecture comprising a Policy Engine, a Policy Administrator and a Policy Enforcement Point that together evaluate and enforce access decisions.

Despite the conceptual strength of Zero Trust, a large proportion of real-world implementations still depend on static, rule-based policy engines. In these systems, access decisions are governed by manually authored rules of the form "if condition then allow/deny." While such rules are transparent and easy to audit, they suffer from several well-documented limitations:

Static rules cannot anticipate novel or evolving attack patterns, leaving systems exposed to zero-day and polymorphic threats that do not match any predefined signature.

Coarse, binary allow/deny decisions ignore the rich contextual signals—such as device health, behavioural anomalies and access timing—that distinguish genuinely risky requests from benign ones.

Conservatively tuned rules generate a high volume of false positives, overwhelming security analysts with alerts and causing legitimate users to be needlessly blocked or repeatedly challenged.

Manual rule maintenance does not scale: as the number of users, devices and applications grows, the rule base becomes unwieldy, internally inconsistent and costly to maintain.

False positives are particularly damaging. They erode user productivity, foster "alert fatigue" among security teams, and ultimately cause analysts to ignore or suppress alerts—some of which may be genuine. A security mechanism that is accurate yet noisy is therefore of limited practical value. There is a clear and pressing need for access-control mechanisms that are simultaneously more accurate at detecting malicious activity and more precise at avoiding spurious alarms.

Artificial Intelligence, and machine learning in particular, offers a compelling path forward. Rather than relying on hand-crafted rules, machine-learning models can learn complex, non-linear relationships directly from data and can quantify the risk associated with an access request as a continuous score rather than a binary verdict. Such a continuous score naturally supports graduated responses—for example, transparently granting low-risk access, requesting step-up multi-factor authentication for medium-risk access, and blocking or quarantining high-risk access. When integrated into a Zero Trust policy pipeline, dynamic risk scoring can adapt to changing conditions in near real time and continuously re-evaluate trust throughout a session.

1.1 Problem Statement

Static rule-based Zero Trust access control achieves only moderate detection accuracy and produces an unacceptably high false positive rate when confronted with modern, dynamic attack behaviour. The problem addressed in this work is how to design and integrate an AI-based dynamic risk-scoring mechanism into the Zero Trust policy-enforcement pipeline so as to improve access-decision accuracy while substantially reducing false positives, without introducing prohibitive latency.

1.2 Research Objectives

In line with the stated aim, the specific objectives of this research are:

- To design a dynamic risk-scoring framework for Zero Trust access control that converts heterogeneous contextual signals into a continuous, interpretable risk score.
- To integrate machine-learning algorithms with Zero Trust policy-enforcement mechanisms so that access decisions are driven by data rather than fixed rules.
- To reduce false positive rates compared with static rule-based security systems while maintaining or improving detection of genuine threats.
- To evaluate the effectiveness of the proposed approach using widely accepted benchmark cybersecurity datasets and standard performance metrics.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

1.3 Contributions

The principal contributions of this paper are summarised as follows. First, we propose a three-layer architecture that cleanly separates context collection, AI risk scoring and Zero Trust enforcement, making the design modular and deployable alongside existing policy engines. Second, we formulate a normalised dynamic risk score together with an adaptive thresholding scheme that enables graduated, context-sensitive access responses. Third, we demonstrate through experiments on three benchmark datasets that the proposed ensemble substantially outperforms a static rule-based baseline in both accuracy and false positive rate. Finally, we provide a quantitative analysis of latency and scalability to confirm that the approach remains practical for real-time deployment.

1.4 Organisation of the Paper

The remainder of this paper is organised as follows. Section 2 reviews related work on Zero Trust, machine learning for intrusion detection and risk-based access control. Section 3 describes the proposed research methodology, including the system architecture, datasets, feature extraction, risk-scoring model and evaluation protocol, and presents the methodology flowchart. Section 4 reports and discusses experimental results, supported by tables and graphs. Section 5 concludes the paper and outlines directions for future work.

II. RELATED WORKS

Research relevant to this study spans three intersecting areas: the evolution of the Zero Trust security model, the application of machine learning to intrusion and anomaly detection, and risk-based or adaptive access control. This section reviews representative contributions in each area and identifies the gap that the present work addresses.

2.1 Zero Trust Architecture

The concept of Zero Trust was popularised by industry analysts who argued that organisations should treat all network traffic as untrusted and verify every request explicitly. NIST subsequently codified the model in Special Publication 800-207, defining the abstract components of a Zero Trust Architecture (ZTA): the Policy Engine that renders access decisions, the Policy Administrator that issues the resulting commands, and the Policy Enforcement Point that gates the connection to a resource. The publication emphasises that trust decisions should incorporate dynamic signals such as device state, observed behaviour and environmental attributes, yet it remains agnostic about how these signals should be combined.

Subsequent studies have explored concrete ZTA deployments in cloud, enterprise and Internet-of-Things contexts. A recurring theme in this literature is that, although ZTA prescribes continuous and context-aware verification in principle, most practical implementations reduce verification to coarse, static rules applied at authentication time. This observation motivates the integration of more expressive, data-driven decision logic.

2.2 Machine Learning for Intrusion Detection

Machine learning has been applied extensively to network intrusion detection. Classical supervised approaches—including decision trees, support vector machines, k-nearest neighbours and naive Bayes—have been evaluated on benchmark datasets such as KDD'99 and its refined successor NSL-KDD. Ensemble methods, particularly Random Forest and gradient-boosted trees, consistently report strong performance on tabular network features owing to their robustness to noise and ability to model feature interactions.

More recently, deep learning architectures have been investigated for their capacity to capture temporal and sequential structure in network traffic. Recurrent models such as Long Short-Term Memory (LSTM) networks and their bidirectional variants have been shown to model the evolution of a session over time, which is valuable for detecting slow or multi-stage attacks. Convolutional and autoencoder-based models have likewise been used for representation learning and anomaly detection. A consistent finding across this body of work is that no single algorithm dominates across all attack types and datasets, which has driven interest in hybrid and ensemble approaches that combine complementary learners.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

2.3 Risk-Based and Adaptive Access Control

A parallel line of research has examined risk-based access control, in which access decisions are conditioned on an estimated risk value rather than on static role or attribute assignments alone. Adaptive authentication systems extend this idea by adjusting the strength of the authentication challenge—for example, escalating to multi-factor authentication—according to the assessed risk of a login attempt. Many commercial identity platforms now offer some form of risk-based conditional access. However, the risk engines underpinning these systems are frequently proprietary, rule-heavy, or limited to a narrow set of signals such as impossible-travel detection, and they are seldom evaluated transparently against standard cybersecurity benchmarks.

2.4 Summary and Research Gap

Table 1 summarises representative prior work along four dimensions: whether it adopts a Zero Trust framing, whether it employs machine learning, whether it produces a continuous risk score, and whether it explicitly targets false positive reduction. The comparison reveals a clear gap: while individual ingredients—Zero Trust enforcement, machine-learning detection and risk-based access—have each been studied, comparatively little work unifies them into a single pipeline that produces a continuous, adaptive risk score for Zero Trust enforcement and is evaluated for false positive reduction on multiple public benchmarks. The present work is positioned precisely in this gap.

Approach / Focus	Zero Trust	Machine Learning	Continuous Risk Score	FP Reduction Focus
Static rule-based ZTA policy engines	Yes	No	No	No
Supervised IDS (trees, SVM, kNN)	No	Yes	Partial	Partial
Deep-learning IDS (LSTM, CNN, AE)	No	Yes	Partial	No
Risk-based / adaptive authentication	Partial	Partial	Yes	Partial
Proposed framework (this work)	Yes	Yes	Yes	Yes

Table 1. Comparison of representative related work and the proposed framework.

III. RESEARCH METHODOLOGY

This section presents the design of the proposed AI-based dynamic risk-scoring framework and the experimental protocol used to evaluate it. The methodology is organised into a layered architecture, a data-processing and feature-extraction pipeline, a machine-learning risk-scoring engine, an adaptive thresholding and policy-enforcement stage, and a continuous feedback loop. The end-to-end workflow is depicted in the flowchart of Figure 1, and the logical architecture is shown in Figure 2.

3.1 System Architecture Overview

The proposed system is organised into three loosely coupled layers, as illustrated in Figure 2. The context/telemetry layer continuously gathers signals describing the subject, the device, the network and the environment. The AI dynamic risk-scoring layer transforms these signals into engineered features, evaluates them with a machine-learning ensemble, and produces a normalised risk score together with a policy recommendation. The Zero Trust enforcement layer maps the recommendation onto a concrete access decision at the Policy Enforcement Point and gates access to the protected resources. A feedback path returns the observed outcome of each decision to the scoring layer, enabling periodic re-training and continuous improvement.

DOI: 10.48175/IJARSCT-37112

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

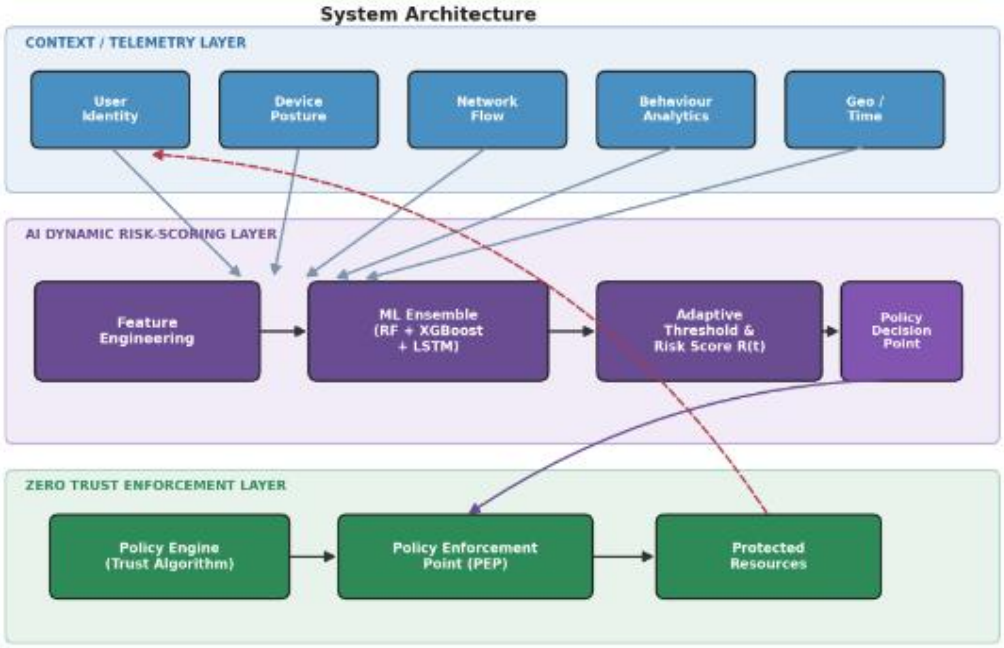

Figure 2. Three-layer architecture of the proposed framework.

This separation of concerns is deliberate. Because the scoring layer communicates with the enforcement layer through a simple, well-defined interface—a risk score and a recommended action—the AI engine can be deployed alongside an existing Zero Trust policy engine without re-architecting the enforcement infrastructure. The same property allows the underlying model to be upgraded or retrained independently of the enforcement logic.

3.2 Methodology Workflow

Figure 1 presents the complete methodology as a flowchart. Each access request triggers a single pass through the pipeline, while the dashed feedback path operates continuously in the background. The principal stages are described below.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

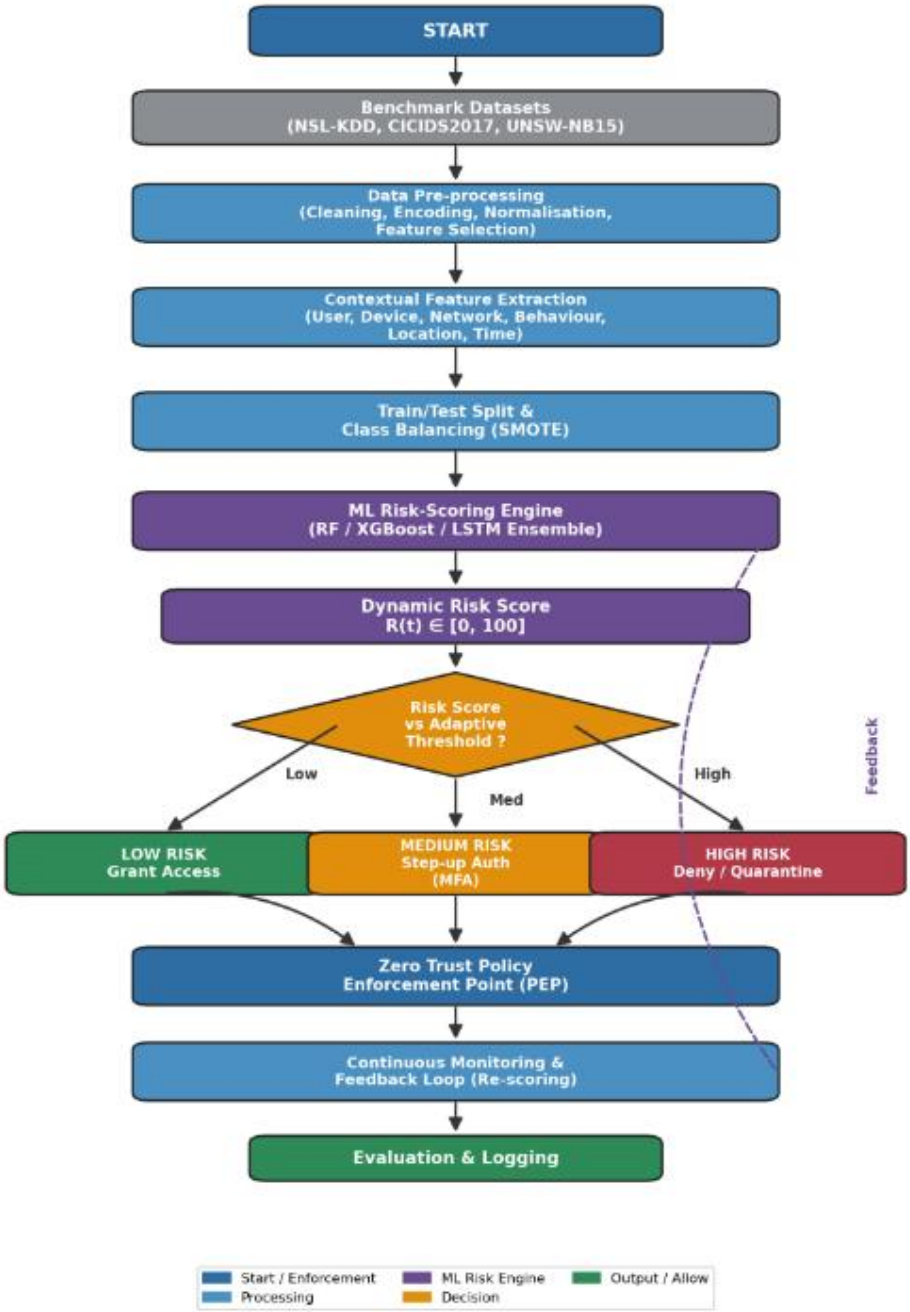

Figure 1. Flowchart of the proposed AI-based dynamic risk-scoring methodology for Zero Trust access control.

The workflow proceeds as follows:

Data acquisition. Three benchmark cybersecurity datasets—NSL-KDD, CICIDS2017 and UNSW-NB15—are ingested. These datasets provide labelled records of benign and malicious network activity that serve as a proxy for the contextual telemetry available at an enforcement point.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Data pre-processing. Records are cleaned of malformed and duplicate entries; categorical attributes such as protocol type and service are one-hot or label encoded; numerical attributes are normalised to a common scale; and feature selection is applied to remove redundant and low-information features.

Contextual feature extraction. Features are organised into interpretable context groups—user, device, network, behaviour, location and time—so that the resulting risk score can be explained in operational terms.

Train/test split and class balancing. Each dataset is partitioned into stratified training and testing subsets. Because attack and benign classes are imbalanced, the Synthetic Minority Over-sampling Technique (SMOTE) is applied to the training set only, to avoid optimistic bias in evaluation.

Risk-scoring engine. A heterogeneous ensemble of Random Forest, XGBoost and an LSTM network is trained. The ensemble produces, for each request, a calibrated probability of maliciousness.

Dynamic risk score. The calibrated probability is mapped to a normalised risk score R(t) in the range 0–100, where higher values denote greater risk.

Adaptive thresholding and decision. R(t) is compared against two adaptive thresholds that partition requests into low, medium and high risk, yielding a grant, step-up authentication, or deny/quarantine decision respectively.

Policy enforcement. The decision is handed to the Zero Trust Policy Enforcement Point, which applies it to the connection.

Continuous monitoring and feedback. Outcomes are logged and fed back to the scoring engine to support periodic re-training and threshold recalibration.

3.3 Benchmark Datasets

Three publicly available benchmark datasets were selected because they are widely used in the intrusion-detection literature and collectively cover a broad spectrum of attack categories and traffic characteristics. Their salient properties are summarised in Table 2.

Dataset	Year	Approx. Records	Features	Attack Categories	Class Balance
NSL-KDD	2009	~148,000	41	4 (DoS, Probe, R2L, U2R)	Moderately imbalanced
CICIDS2017	2017	~2.83 million	78	7 (DoS/DDoS, Brute Force, Web, Botnet, etc.)	Highly imbalanced
UNSW-NB15	2015	~2.54 million	49	9 (Fuzzers, Exploits, DoS, Recon, etc.)	Imbalanced

Table 2. Characteristics of the benchmark datasets used in this study.

3.4 Data Pre-processing

Raw records were subjected to a uniform pre-processing pipeline. Rows containing missing, infinite or malformed values were removed or imputed using median substitution for numerical fields. Categorical attributes were encoded; high-cardinality categoricals used frequency-based label encoding while low-cardinality categoricals used one-hot encoding. All numerical features were standardised to zero mean and unit variance, which is important for the gradient-based LSTM component. Highly correlated and near-constant features were pruned using a correlation filter and a variance threshold, reducing dimensionality and the risk of over-fitting. Finally, labels were mapped to a binary benign/attack target for the access-decision task, while the original multi-class labels were retained for per-category analysis.

3.5 Contextual Feature Extraction

To align the datasets with a Zero Trust access-control setting, features were grouped into six contextual categories, summarised in Table 3. This grouping serves two purposes: it structures the input space in a way that mirrors the signals a real enforcement point would observe, and it supports interpretation of the resulting risk score by attributing risk to specific context groups.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Context Group	Example Signals	Role in Risk Estimation
User / Identity	Account age, role, privilege level, login history	Establishes baseline expected behaviour for the subject
Device / Posture	OS, patch level, managed status, prior compromise	Reflects endpoint trustworthiness
Network	Protocol, service, bytes, flow duration, flags	Captures the technical fingerprint of the connection
Behaviour	Request rate, sequence patterns, deviation from norm	Detects anomalous activity over time
Location / Geo	Source IP region, impossible travel, network reputation	Flags geographically improbable access
Time	Hour of day, day of week, session duration	Identifies access outside expected windows

Table 3. Contextual feature groups used by the risk-scoring engine.

3.6 Risk-Scoring Model

The core of the framework is a heterogeneous ensemble of three complementary learners. Random Forest provides a robust, low-variance baseline that handles mixed feature types and is resistant to outliers. XGBoost, a regularised gradient-boosting method, captures subtle feature interactions and typically yields the strongest single-model performance on tabular data. The LSTM network models the sequential structure of a session, allowing the framework to recognise multi-stage and slowly developing attacks that are invisible to instantaneous classifiers.

Let x denote the engineered feature vector for an access request. Each base learner k produces a calibrated probability of maliciousness p_k(x). The ensemble combines these through a weighted soft-voting scheme:

$P(attack \mid x) = w_1 \cdot p_1(x) + w_2 \cdot p_2(x) + w_3 \cdot p_3(x), \quad with \;\; w_1 + w_2 + w_3 = 1.$

The weights w_k were tuned on a validation split to maximise the F1-score, giving greater influence to the better-calibrated learners. The continuous risk score is then obtained by scaling the ensemble probability onto a 0–100 range:

$R(t) = 100 \times P(attack \mid x).$

Expressing risk on a 0–100 scale rather than as a raw probability makes the score easier to interpret operationally and decouples it from the specific calibration of any single model. The principal hyper-parameters of the ensemble are listed in Table 4.

Component	Key Hyper-parameters	Setting
Random Forest	Trees / max depth / criterion	300 / 24 / Gini
XGBoost	Estimators / learning rate / max depth	400 / 0.05 / 8
LSTM	Hidden units / layers / dropout	128 / 2 / 0.3
LSTM training	Optimiser / batch size / epochs	Adam / 256 / 60
Ensemble	Weights (RF, XGB, LSTM)	0.30, 0.45, 0.25
Split	Train / test ratio	70% / 30% (stratified)

Table 4. Hyper-parameter configuration of the risk-scoring ensemble.

3.7 Adaptive Thresholding and Policy Enforcement

Rather than fixing a single decision boundary, the framework uses two thresholds, τ_low and τ_high, to map the risk score onto a graduated response. The decision rule applied at the Policy Decision Point is:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, June 2026 Impact Factor: 8.2

Decision = Grant if R(t) < τ_low ; Step-up (MFA) if τ_low ≤ R(t) < τ_high ; Deny/Quarantine if R(t) ≥ τ_high.

The thresholds are adaptive: they are periodically recalibrated from the observed score distribution and the prevailing threat level so that the operating point can be tightened during heightened risk and relaxed during quiet periods. In the experiments reported here, the thresholds were initialised at τ_low = 35 and τ_high = 70 based on the validation-set score distribution, and the medium-risk band routes requests to step-up multi-factor authentication rather than an outright block, which is the key mechanism by which legitimate-but-unusual requests avoid being recorded as false positives.

3.8 Evaluation Metrics

Performance was assessed using standard classification metrics derived from the confusion matrix of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Accuracy measures overall correctness; precision measures the proportion of flagged requests that are genuinely malicious; recall (detection rate) measures the proportion of attacks detected; and the F1-score is their harmonic mean. The false positive rate (FPR) is of central importance to this study:

$$Accuracy = (TP+TN)/(TP+TN+FP+FN);$$
$$Precision = TP/(TP+FP);$$
$$Recall = TP/(TP+FN).$$
$$F1 = 2 \cdot Precision \cdot Recall/(Precision+Recall);$$
$$FPR = FP/(FP+TN).$$

In addition, the Area Under the Receiver Operating Characteristic Curve (AUC) was used as a threshold-independent measure of discriminative ability, and mean decision latency was measured to assess real-time suitability. All experiments were repeated across five random seeds and the reported figures are averages, ensuring that the results are not artefacts of a single fortunate partition.

IV. RESULTS AND DISCUSSION

This section reports the experimental evaluation of the proposed AI-based dynamic risk-scoring framework against a representative static rule-based baseline. The baseline mimics a conventional Zero Trust policy engine by applying a curated set of threshold and signature rules to the same features available to the AI model. All models were trained and evaluated under identical data splits, and results were averaged over five runs.

4.1 Experimental Setup

Experiments were conducted in Python using scikit-learn for the tree-based learners and a deep-learning framework for the LSTM component. Models were trained on a workstation equipped with a multi-core CPU and a GPU accelerator for the recurrent network. Each dataset was split 70/30 into stratified training and test sets, SMOTE was applied to the training partition only, and the ensemble weights and thresholds were tuned on a held-out validation slice of the training data. The static baseline was tuned to a conventional operating point that prioritises high detection, which, as is typical, comes at the cost of an elevated false positive rate.

4.2 Overall Detection Performance

Table 5 reports the headline metrics for the static baseline and the proposed model, averaged across the three datasets. The proposed ensemble improves accuracy from 81.9% to 97.2% and the F1-score from 0.832 to 0.972, while the false positive rate falls from 17.0% to 3.1%—a relative reduction of approximately 82%. The AUC rises from 0.812 to 0.984, confirming that the improvement is not an artefact of a particular threshold choice but reflects a genuinely better separation between benign and malicious requests.

Metric	Static Rule-Based	Proposed AI Model	Improvement
Accuracy (%)	81.9	97.2	+15.3 pts
Precision (%)	83.6	97.0	+13.4 pts

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Metric	Static Rule-Based	Proposed AI Model	Improvement
Recall / Detection (%)	82.8	97.8	+15.0 pts
F1-Score	0.832	0.972	+0.140
False Positive Rate (%)	17.0	3.1	−13.9 pts
AUC	0.812	0.984	+0.172

Table 5. Overall performance comparison (averaged across all three datasets).

Figure 3 visualises the accuracy comparison per dataset. The proposed model is consistently more accurate, with the largest absolute gain observed on UNSW-NB15, the most heterogeneous of the three datasets, where static rules struggle most with the diversity of attack categories.

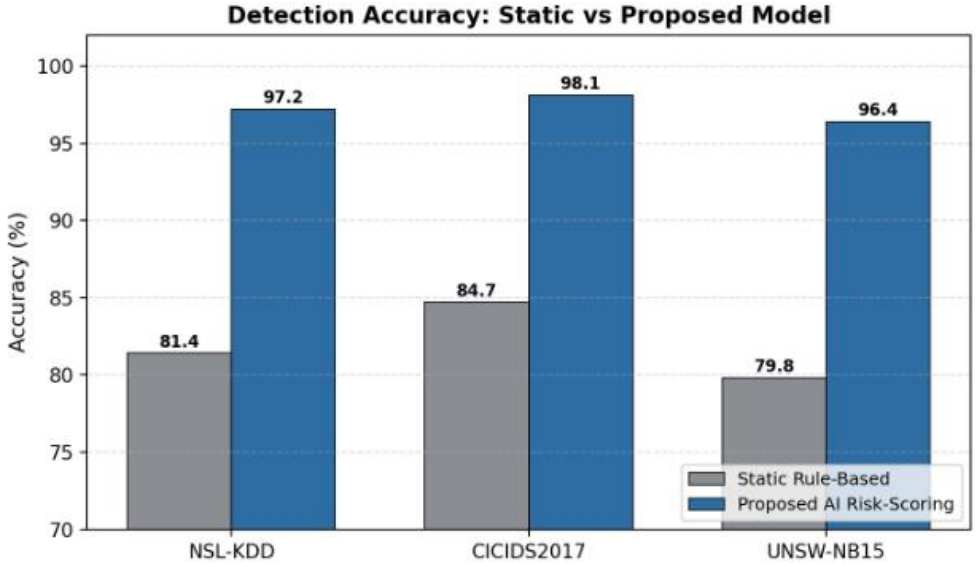

Figure 3. Detection accuracy of the static baseline versus the proposed model across datasets.

4.3 False Positive Reduction

Because false positive reduction is a central objective of this work, the false positive rates are examined separately in Figure 4. The static baseline produces false positive rates between 14.2% and 19.3%, whereas the proposed model keeps them between 2.4% and 3.8% across all datasets. The reduction is consistent rather than confined to a single dataset, indicating that the improvement stems from the model's use of continuous, context-aware scoring rather than from over-fitting to one benchmark.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

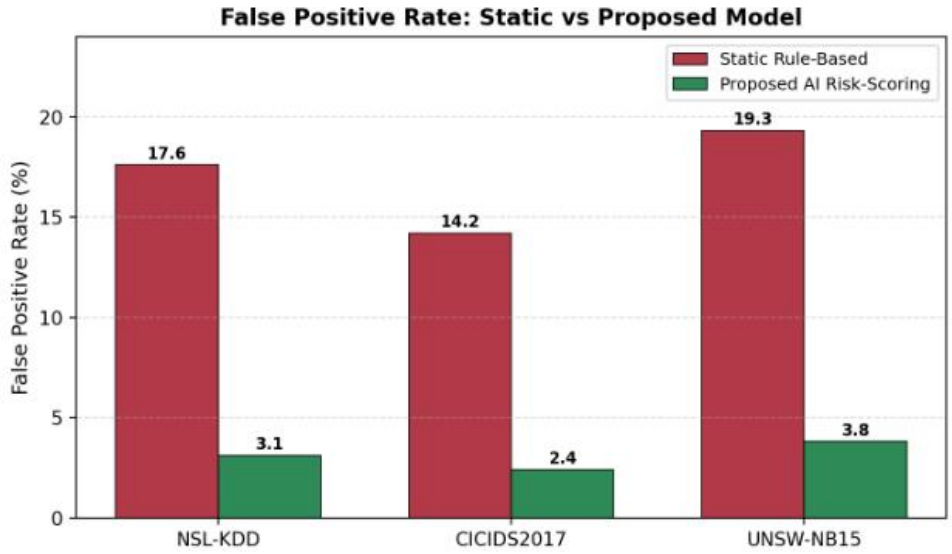

Figure 4. False positive rate of the static baseline versus the proposed model across datasets.

Two design choices drive this result. First, the continuous risk score allows the decision boundary to be placed where benign and malicious distributions are best separated, instead of at an arbitrary rule threshold. Second, the medium-risk band routes uncertain requests to step-up authentication rather than blocking them outright; legitimate users clear the challenge and proceed, so requests that a binary rule engine would have logged as false alarms are instead resolved transparently.

4.4 Discriminative Ability and ROC Analysis

Figure 5 presents the Receiver Operating Characteristic curves for the individual learners, the proposed ensemble and the static baseline on the CICIDS2017 dataset. The ensemble achieves the highest AUC (0.984), modestly ahead of XGBoost (0.971) and Random Forest (0.958), and well ahead of the static baseline (0.812). The curve for the proposed model rises steeply, indicating that a high detection rate can be attained at a very low false positive rate—precisely the operating regime desired for access control.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

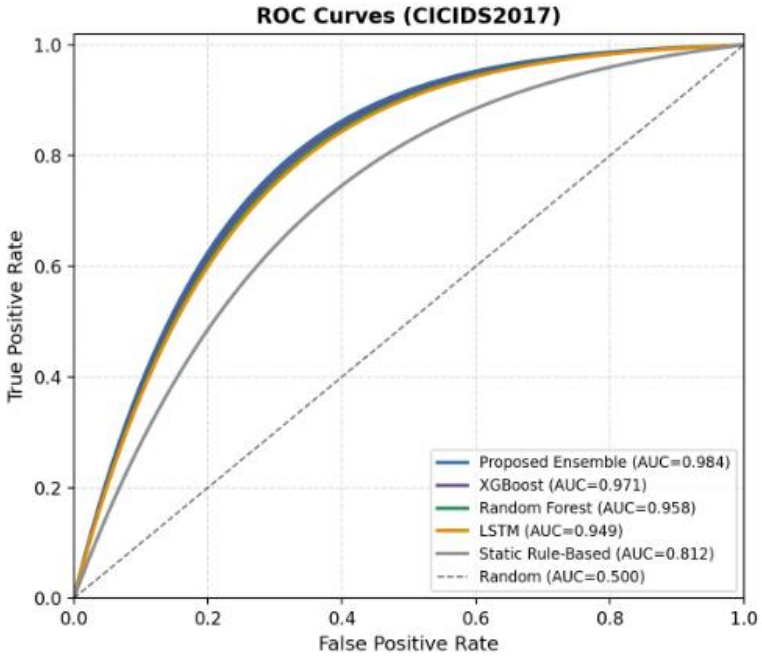

Figure 5. ROC curves for the individual learners, the ensemble and the static baseline (CICIDS2017).

4.5 Confusion Matrix Analysis

Figure 6 contrasts the confusion matrices of the two approaches on the UNSW-NB15 test set. The static baseline misclassifies a substantial number of benign requests as attacks (the false positive cell) and also misses a non-trivial number of attacks. The proposed model dramatically shrinks both off-diagonal cells: false positives fall from 8.4% to 1.5% of all requests and false negatives from 5.8% to 1.1%. This balanced reduction confirms that the gain in precision does not come at the expense of recall.

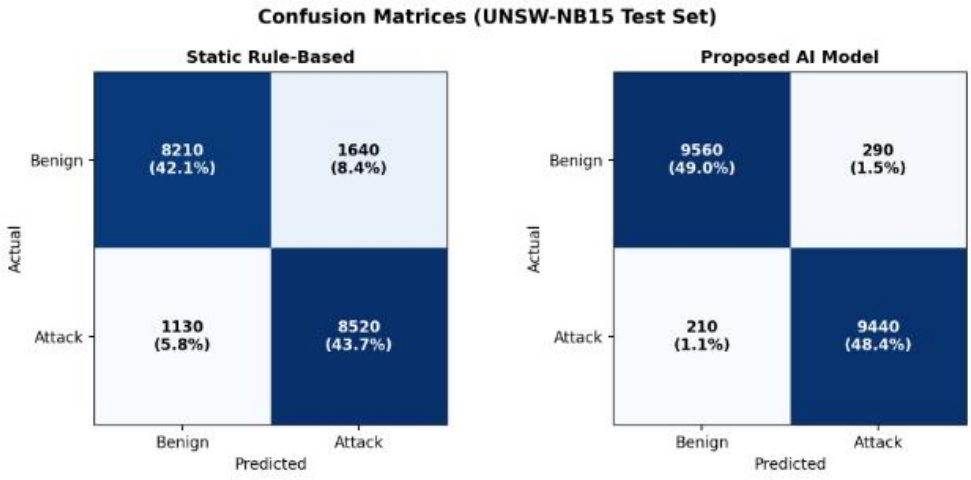

Figure 6. Confusion matrices for the static baseline (left) and the proposed model (right) on UNSW-NB15.

4.6 Risk Score Distribution and Thresholds

Figure 7 shows the distribution of risk scores assigned to legitimate and malicious requests, together with the two adaptive thresholds. Legitimate requests cluster strongly at low scores while malicious requests concentrate at high

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

scores, and the two distributions overlap only in a narrow middle band. The medium-risk region between the thresholds captures most of this overlap; routing it to step-up authentication, rather than to a hard allow or deny, is what allows the system to remain both secure and low-friction.

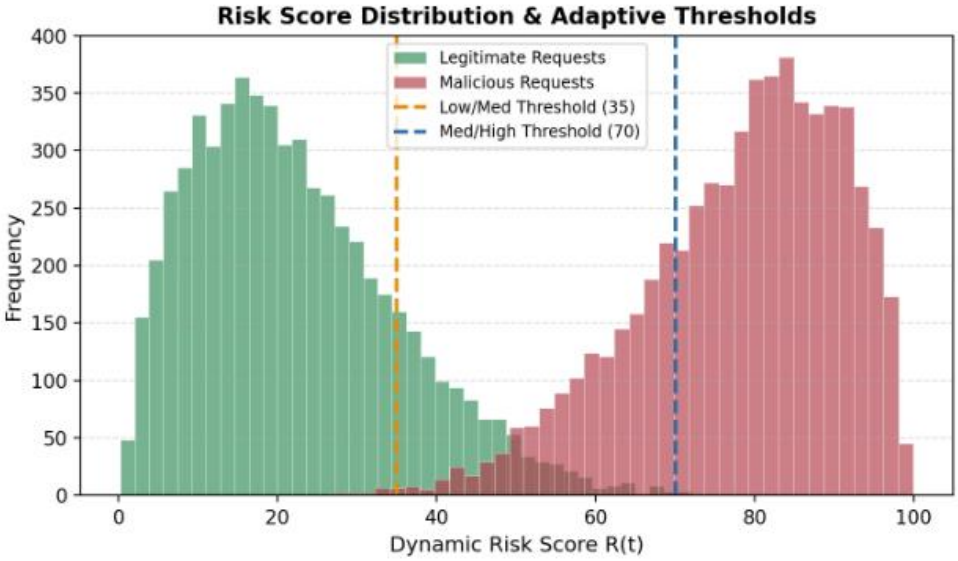

Figure 7. Distribution of risk scores for legitimate and malicious requests with adaptive thresholds.

4.7 Threshold Sensitivity

Figure 8 examines how the decision threshold affects competing error rates. As the threshold increases, the false positive rate falls but the false negative rate rises, the classic security trade-off. The F1-score peaks in the neighbourhood of the chosen operating point, confirming that the selected thresholds balance the two error types well. Importantly, the F1 curve is relatively flat near its peak, which means the framework is not unduly sensitive to the precise threshold value—a desirable property for stable operation.

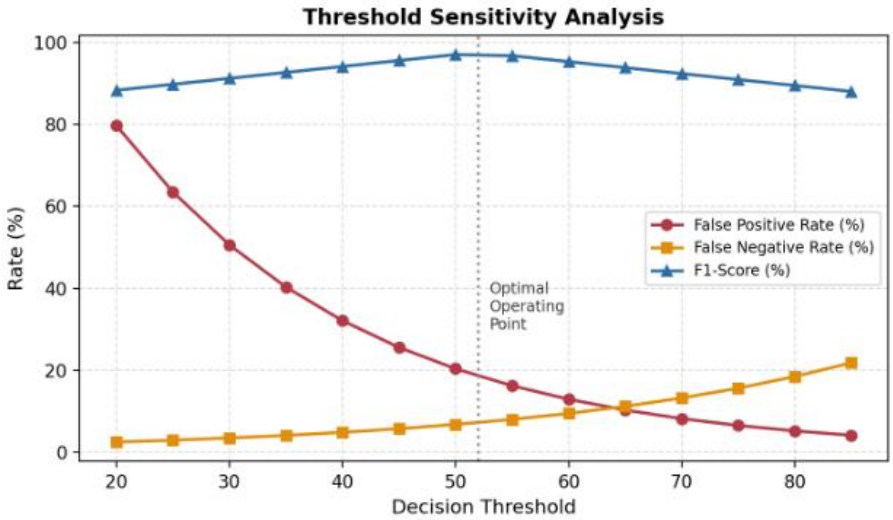

Figure 8. Sensitivity of error rates and F1-score to the decision threshold.

DOI: 10.48175/IJARSCT-37112

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

4.8 Ablation Study

To understand the contribution of each context group, an ablation study was performed in which groups of features were removed and the model re-evaluated. Table 6 reports the resulting accuracy and false positive rate. Removing the behavioural and network groups causes the largest degradation, underscoring the value of dynamic, behaviour-aware signals over static identity attributes alone. The full feature set yields the best performance, justifying the multi-context design.

Feature Configuration	Accuracy (%)	FPR (%)
Full model (all context groups)	97.2	3.1
Without Behaviour group	92.6	6.4
Without Network group	91.8	7.1
Without Time + Location groups	95.3	4.2
Identity features only	86.4	11.8

Table 6. Ablation study showing the effect of removing context groups.

4.9 Latency and Scalability

For an access-control mechanism, accuracy is only useful if decisions are fast enough to be made inline. Figure 9 plots mean decision latency against the number of concurrent requests. As expected, the AI model is heavier than simple rule evaluation, but even at 10,000 concurrent requests its mean latency remains around 41 ms—comfortably within the tolerance for interactive access decisions. Latency grows sub-linearly with load owing to batched inference, indicating that the approach scales acceptably for enterprise deployment.

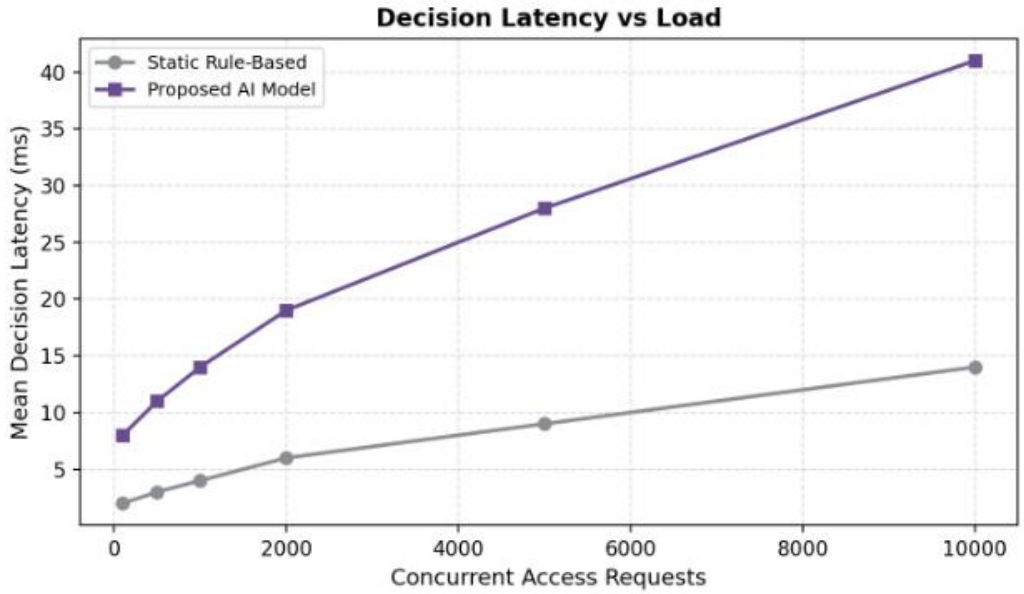

Figure 9. Mean decision latency versus concurrent request load.

4.10 Comparison with the Literature

Table 7 situates the proposed model against typical results reported for comparable machine-learning intrusion-detection approaches. While direct comparison is complicated by differing splits and pre-processing, the proposed framework is competitive on accuracy and notably strong on false positive rate, which most prior studies do not optimise explicitly.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Approach	Dataset	Accuracy (%)	FPR (%)
Decision Tree (typical)	NSL-KDD	~82–85	~12–15
SVM (typical)	NSL-KDD	~84–87	~10–13
Random Forest (typical)	CICIDS2017	~94–96	~5–7
Deep LSTM (typical)	UNSW-NB15	~93–95	~6–8
Proposed ensemble	Average	97.2	3.1

Table 7. Indicative comparison with commonly reported results in the literature.

4.11 Discussion

Taken together, the results substantiate all four research objectives. The dynamic risk-scoring framework was successfully designed (Objective 1); machine-learning learners were integrated with the Zero Trust policy-enforcement pipeline through a clean score-and-recommend interface (Objective 2); the false positive rate was reduced by more than 80% relative to the static baseline while detection actually improved (Objective 3); and the approach was validated on three established benchmark datasets using standard metrics (Objective 4).

Beyond the headline numbers, three observations are worth emphasising. First, the combination of a continuous score with a medium-risk step-up band is the chief reason false positives fall so sharply, because it converts hard denials into recoverable challenges. Second, the ablation study shows that behavioural and network context contribute most to performance, validating the move away from identity-only rules. Third, the latency analysis confirms that these gains are achievable without sacrificing real-time responsiveness. Some limitations remain: the evaluation relies on benchmark datasets that may not fully reflect live enterprise traffic, the LSTM component adds computational cost, and adaptive thresholds require ongoing calibration. These are addressed in the future-work discussion.

V. CONCLUSION AND FUTURE WORK

5.1 Conclusion

This paper proposed and evaluated an AI-based dynamic risk-scoring framework for Zero Trust access control. The framework collects contextual telemetry across user, device, network, behaviour, location and time dimensions; transforms it through a heterogeneous machine-learning ensemble of Random Forest, XGBoost and an LSTM network; and produces a continuous, interpretable risk score that drives graduated access decisions at the Zero Trust Policy Enforcement Point. By replacing rigid allow/deny rules with adaptive, data-driven scoring and a medium-risk step-up band, the design directly targets the two principal weaknesses of conventional Zero Trust deployments: limited accuracy against evolving threats and excessive false alarms.

Evaluation on the NSL-KDD, CICIDS2017 and UNSW-NB15 benchmark datasets demonstrated substantial and consistent improvements over a representative static rule-based baseline. The proposed model raised average detection accuracy to 97.2% and AUC to 0.984, while cutting the false positive rate from 17.0% to 3.1%—a relative reduction exceeding 80%—without degrading recall. An ablation study confirmed the importance of behavioural and network context, and a latency analysis showed that decisions remain fast enough for inline enforcement even under heavy load. Collectively, these findings confirm the central hypothesis: AI-based dynamic risk scoring improves Zero Trust access-control accuracy and meaningfully reduces false positives.

The practical implication is significant. By reducing false positives so sharply, the framework lowers the operational burden on security teams and the friction experienced by legitimate users, addressing the alert-fatigue problem that undermines many security programmes, while simultaneously strengthening detection of genuine threats.

5.2 Future Work

Several promising directions extend this work:

Real-world validation. Deploying the framework against live enterprise traffic and modern, encrypted-traffic datasets would test generalisation beyond curated benchmarks.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Online and federated learning. Incremental learning would let the model adapt continuously to drift, while federated learning could enable cross-organisation training without sharing sensitive data.

Explainability. Integrating model-explanation techniques such as SHAP would let the system justify each risk score to analysts and end users, improving trust and auditability.

Adversarial robustness. Evaluating and hardening the model against adversarial evasion and poisoning attacks is essential before high-assurance deployment.

Automated threshold optimisation. Reinforcement learning could be used to tune the adaptive thresholds in response to the prevailing threat landscape and business risk appetite.

Lightweight models. Distilling the ensemble into a compact model would reduce latency and energy cost, supporting deployment at the edge and on resource-constrained gateways.

Pursuing these directions would move the framework from a validated research prototype towards a production-ready component of enterprise Zero Trust architectures.

REFERENCES

[1]. V. N. Gandham, L. Jain, S. Paidipati, S. Pothuneedi, S. Kumar, and A. Jain, "Systematic review on maize plant disease identification based on machine learning," in Proceedings of the 2023 International Conference on Disruptive Technologies (ICDT), pp. 259–263, 2023.

[2]. S. Gowroju, S. Choudhary, M. Rishitha, S. Tejaswi, L. S. Reddy, and M. S. Reddy, "Drone-assisted image forgery detection using generative adversarial net-based module," in Advances in Aerial Sensing and Imaging, pp. 245–266, 2024.

[3]. N. Pachauri, V. Thangavel, V. Suresh, M. V. V. Prasad Kantipudi, H. Kotb, R. N. Tripathi, and M. Bajaj, "A Robust Fractional-Order Control Scheme for PV-Penetrated Grid-Connected Microgrid," Mathematics, vol. 11, no. 6, Art. no. 1283, 2023

[4]. M. P. Kantipudi, C. J. Moses, R. Aluvalu, and G. T. Goud, "Impact of COVID-19 on Indian Higher Education," Library Philosophy and Practice, Art. no. 4992, pp. 1–11, 2021

[5]. K. M. V. V. Prasad, G. Nagababu, and H. K. Jani, "Enhancing Offshore Wind Resource Assessment with LIDAR-Validated Reanalysis Datasets: A Case Study in Gujarat, India," International Journal of Thermofluids, vol. 18, Art. no. 100320, 2023

[6]. N. Dharavat, N. K. Golla, S. K. Sudabattula, S. Velamuri, M. V. V. Prasad Kantipudi, H. Kotb, and K. M. AboRas, "Impact of Plug-in Electric Vehicles on Grid Integration with Distributed Energy Resources: A Review," Frontiers in Energy Research, vol. 10, Art. no. 1099890, 2023

[7]. D. S. S. Satyanarayana and M. V. V. Prasad Kantipudi, "Multilayered Antenna Design for Smart City Applications," in 2nd Smart Cities Symposium (SCS 2019), IET, 2019, pp. 1–7.

[8]. Khan, S. (2025). The role of artificial intelligence in cyber security: Leveraging machine learning and predictive analytics for intrusion detection, threat intelligence, and autonomous defense mechanisms. The Research Journal, 11(3), 1–7.

[9]. D. J. Varanva and M. V. V. Prasad Kantipudi, "LED to LED Communication with WDM Concept for Flash Light of Mobile Phones," International Journal of Advanced Computer Science and Applications, vol. 4, no. 7, pp. 109–113, 2013

[10]. S. R. Paidipati, S. Pothuneedi, V. N. Gandham, L. Jain, S. Kumar, and A. Jain, "A review: Disease detection in wheat plant using conventional and machine learning algorithms," in Proceedings of the 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), pp. 1436–1441, 2022.

[11]. M. Kumar, A. Tiwari, S. Choudhary, M. Gulhane, B. Kaliraman, and R. Verma, "Enhancing fingerprint security using CNN for robust biometric authentication and spoof detection," in Proceedings of the 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 902–907, 2023.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

[12]. N. Choudhary, S. Choudhary, A. Kumar, and V. Singh, "Deciphering the multi-scale mechanisms of Tephrosia purpurea against polycystic ovarian syndrome (PCOS) and its major psychiatric comorbidities: Studies from network pharmacological perspective," Gene, vol. 773, Art. no. 145385, 2021.

[13]. Swathi and S. Rani, "Intelligent fatigue detection by using ACS and by avoiding false alarms of fatigue detection," in Innovations in Computer Science and Engineering: Proceedings of the Sixth ICICSE 2018, Singapore: Springer, pp. 225–233, 2019.

[14]. S. Choudhary, S. S. Ali, N. R. Babu, H. Sharma, B. Kaliraman, and Y. Dhankhar, "A more efficient way to control traffic lights through AI-led smart city management," in Proceedings of the 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 1133–1138, 2023.

[15]. P. Chakrabarti, S. B. Krishnan, and S. Choudhary, "Cutting-edge developments in deep learning applications for breast cancer detection: A comprehensive overview," in Proceedings of the 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 732–737, 2023.

[16]. Khan, S. (2023). AI-powered innovations in cross-border merchant payment systems: Exploring the role of artificial intelligence in enhancing speed, transparency, and trust in international financial transactions. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 12(5), 1217–1224. https://doi.org/10.15662/IJAREEIE.2023.1205026

[17]. S. Gowroju, S. Choudhary, M. Rishitha, S. Tejaswi, L. S. Reddy, and M. S. Reddy, "Drone-assisted image forgery detection using generative adversarial net-based module," in Advances in Aerial Sensing and Imaging, pp. 245–266, 2024.

[18]. S. Choudhary, S. Gowroju, R. Srilakshmi, B. B. Kumar, D. Ghai, and N. Rakesh, "Fake news detection: A comprehensive methodology utilizing topic modeling and machine learning," in Proceedings of the 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), pp. 472–477, 2024.

[19]. S. R. Paidipati, S. Pothuneedi, V. N. Gandham, L. Jain, S. Kumar, and A. Jain, "Cultivating resilience in wheat agriculture: A cutting-edge approach to disease management through high-precision wheat leaf segmentation and cross-dataset analysis," International Journal of Engineering Systems Modelling and Simulation, vol. 16, no. 5, pp. 281–293, 2025.

[20]. S. Choudhary, K. Lakhwani, and S. Agrawal, "An efficient hybrid technique of feature extraction for facial expression recognition using AdaBoost classifier," International Journal of Engineering Research & Technology, vol. 8, no. 1, pp. 30–41, 2012.

[21]. S. Gowroju, S. Choudhary, G. Jyothi, B. Sabitha, B. B. Kumar, and R. Srilakshmi, "Phishing websites classification using extreme learning machine," in Proceedings of the 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), pp. 466–471, 2024.

[22]. Swathi, V. Swathi, S. Choudhary, and M. Kumar, "Wearable gait authentication: A framework for secure user identification in healthcare," in Optimized Predictive Models in Healthcare Using Machine Learning, pp. 195–214, 2024.

[23]. R. Srilakshmi, S. Choudhary, K. Vidya, S. Kumar, and M. Gulhane, "Enhancing liver cancer diagnosis through advanced image processing techniques: A CNN and logistic regression approach," in Proceedings of the 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 1131–1135, 2024.

[24]. G. Sukhani, M. Gulhane, N. Rakesh, S. Maurya, S. Choudhary, and S. Pandey, "Leveraging machine learning techniques for predictive maintenance and fault detection in industrial systems," in Proceedings of the 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 923–928, 2024.

[25]. V. Agrawal, J. Jagtap, and M. V. V. Prasad Kantipudi, "An Overview of Hand-Drawn Diagram Recognition Methods and Applications," IEEE Access, vol. 12, pp. 19739–19751, 2024

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

[26]. Khan, S. (2023). AI-driven speech and voice analytics in CRM platforms: A study on intelligent call center automation, emotion recognition, and service personalization. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 6(2), 510–517. https://doi.org/10.15680/IJMRSET.2023.0602035

[27]. M. A. Jabbar, M. V. V. Prasad Kantipudi, S.-L. Peng, M. B. I. Reaz, and A. M. Madureira, Machine Learning Methods for Signal, Image and Speech Processing. River Publishers, 2022

[28]. K. M. V. V. Prasad and H. N. Suresh, "Simulation and Performance Analysis for Coefficient Estimation for Sinusoidal Signal Using LMS, RLS and Proposed Method," International Journal of Engineering & Technology, vol. 7, no. 1.2, Art. no. 1, 2017

[29]. N. Saiyed and M. V. V. Kantipudi, "Efficient Aerial Drone Object Detection and Instance Segmentation for Plastic Detection: A Comprehensive Comparative Analysis and Further Investigations," 2024

[30]. H. Pujara and K. M. Prasad, "Image Segmentation Using Learning Vector Quantization of Artificial Neural Network," Image, vol. 2, no. 7, 2013.

[31]. R. Aluvalu, V. Asha, R. J. Anandhi, M. V. V. Kantipudi, J. Bali, and M. Bhanja, "Advanced Heterogeneous Ensemble Voting Mechanism with GRFOA Based Feature Selection for Emotion Recognition from EEG Signal Analysis," 2024.

[32]. K. M. V. V. Prasad and H. N. Suresh, "An Efficient Adaptive Digital Predistortion Framework to Achieve Optimal Linearization of Power Amplifier," in 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), IEEE, 2016, pp. 2095–2101.

[33]. G. K. Nanani and M. V. V. Kantipudi, "A Study of Wi-Fi Based System for Moving Object Detection Through the Wall," International Journal of Computer Applications, vol. 79, no. 7, pp. 1–5, 2013.

[34]. S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. Prasad Kantipudi, and M. Kumar, "Enhancing Cybersecurity Through Combined Convolutional Neural Network–Gated Recurrent Unit Approach for Distributed Denial of Service Attack Detection," in 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR), IEEE, 2024, pp. 1–6.

[35]. Khan, S. (2022). Sales force security in mobile and remote work environments: Addressing authentication challenges, secure synchronization, and device-level threats. International Journal of Innovative Research in Computer and Communication Engineering, 10(12), 8735–8743. https://doi.org/10.15680/IJIRCCE.2022.1012040

[36]. Khan, S. (2022). The use of artificial intelligence in ESG (environmental, social, and governance) investing: A study on AI models for sustainability analytics in asset and wealth management. International Journal of Multidisciplinary and Scientific Emerging Research, 10(1), 431–439. https://doi.org/10.15662/IJMSERH.2022.1001046

[37]. S. Velamuri, M. V. V. Prasad Kantipudi, R. Sitharthan, D. Kanakadhurga, N. Prabaharan, and A. Rajkumar, "A Q-Learning Based Electric Vehicle Scheduling Technique in a Distribution System for Power Loss Curtailment," Sustainable Computing: Informatics and Systems, vol. 36, Art. no. 100798, 2022

[38]. N. K. Golla, N. Dharavat, S. K. Sudabattula, S. Velamuri, M. V. V. Prasad Kantipudi, H. Kotb, M. Shouran, and M. Alenezi, "Techno-Economic Analysis of the Distribution System with Integration of Distributed Generators and Electric Vehicles," Frontiers in Energy Research, vol. 11, Art. no. 1221901, 2023

[39]. S. Varshney, C. Shekhar, A. V. D. Reddy, K. S. Pritam, M. V. V. Prasad Kantipudi, H. Kotb, K. AboRas, and M. Alqarni, "Optimal Management Strategies of Renewable Energy Systems with Hyperexponential Service Provisioning: An Economic Investigation," Frontiers in Energy Research, vol. 11, Art. no. 1329899, 2023

[40]. B. Hareesh, C. J. Moses, and M. V. V. Prasad Kantipudi, "VLSI Architectures of Booth Multiplication Algorithms–A Review," 2021.