

RESEARCH ARTICLE

A Hybrid Multi-Layer Framework for Real-Time Malware and Phishing Detection Integrating Logistic Regression with External Threat Intelligence Validation

Rumman Firdos¹, Aman Dangi²

ABSTRACT: The escalating sophistication of malware variants and phishing campaigns has created an urgent need for detection systems that balance high accuracy with practical deployability under real-world constraints including low latency, limited computational resources, and noisy threat intelligence feeds. While deep learning approaches have demonstrated impressive detection rates, they frequently suffer from limited interpretability, elevated false-positive rates, and prohibitive operational costs that hinder their adoption in production security environments. This paper introduces SecureScan, a calibrated multi-layer threat detection framework that formally integrates heuristic analysis, lightweight machine learning, and external threat intelligence through a probabilistic consensus model. Unlike existing hybrid systems that rely on sequential or heuristic decision rules, SecureScan implements a calibrated gray-zone decision mechanism that explicitly quantifies classification uncertainty and dynamically optimizes decision thresholds to minimize false positives while maintaining detection sensitivity. The proposed framework employs logistic regression as an interpretable statistical foundation, enhanced by mathematically grounded confidence fusion and verification layers that enable robust performance without depending on complex deep learning architectures. Comprehensive experiments conducted on a large-scale dataset comprising over 800,000 benign and malicious samples demonstrate that SecureScan achieves 93.1% accuracy, 0.87 precision, and 0.92 recall, significantly reducing false-positive rates and inference latency compared to representative state-of-the-art machine learning and intelligence-driven baselines. These findings suggest that carefully designed, calibrated lightweight models can outperform more complex approaches in operational cybersecurity environments, offering a practical and scientifically validated alternative for real-time malware and phishing detection.

Keywords: Malware Detection; Phishing Detection; Threat Intelligence Integration; Logistic Regression; Uncertainty Calibration; Lightweight Machine Learning for Cybersecurity

Received: 23 January 2026; Revised: 20 February 2026; Accepted: 29 March 2026; Published Online: 27 April 2026

1. INTRODUCTION

The contemporary cybersecurity landscape faces

¹ School of Engineering Science and Technology, Jamia Hamdard University, Hamdard Nagar, New Delhi-110062, Delhi, India.

² Madhav Institute of Technology & Science (MITS), Gola ka Mandir, Gwalior - 474005, Madhya Pradesh, India

* Author to whom correspondence should be addressed:
firdosrumman@gmail.com (Rumman Firdos)

unprecedented challenges as threat actors continuously refine their methodologies to evade traditional defensive measures. The proliferation of polymorphic malware variants, increasingly sophisticated phishing campaigns, and the emergence of advanced persistent threats have collectively rendered conventional security approaches increasingly inadequate for protecting modern digital infrastructures [1]. Traditional intrusion detection systems and antivirus solutions that depend exclusively on signature-based

detection mechanisms operate on a fundamental limitation, they can only identify threats that have been previously observed, documented, and incorporated into signature databases. This reactive paradigm leaves organizations vulnerable to zero-day exploits and novel attack vectors that exploit previously unknown vulnerabilities or employ obfuscation techniques specifically designed to circumvent signature-based detection [2].

The inherent limitations of signature-based approaches have motivated extensive research into machine learning methodologies for cybersecurity applications. Machine learning offers the potential for adaptive, data-driven threat detection that can identify malicious patterns without requiring explicit signatures for each individual threat variant [3]. Statistical learning algorithms can be trained on large datasets comprising both benign and malicious samples to automatically discover discriminative features that distinguish between safe and harmful content. This capability is particularly valuable for detecting phishing websites, which frequently employ URL structures and content layouts that deviate from legitimate patterns in subtle but detectable ways [4]. Similarly, malware detection benefits from machine learning approaches that can identify malicious executables based on static features extracted from file headers, byte sequences, or imported library functions without requiring dynamic execution analysis [5].

Despite these advantages, purely machine learning-based detection systems introduce their own set of significant challenges that limit their practical deployment in operational security environments. The first challenge concerns explainability, complex models such as deep neural networks operate as black boxes that produce classification decisions without providing interpretable rationales for their outputs [6]. This opacity creates difficulties for security analysts who need to understand why particular files or URLs were flagged as malicious, which is essential for incident response, forensic investigation, and regulatory compliance. The second challenge involves false-positive rates, machine learning models frequently exhibit elevated false-positive rates on borderline cases where classification uncertainty is high, leading to alert fatigue and reduced analyst productivity [7]. These false alarms can be particularly problematic in enterprise environments where legitimate traffic exhibits unusual patterns due to content delivery networks, URL shorteners, or corporate redirect mechanisms. The third challenge concerns sensitivity to training data drift, machine learning models trained on historical data may experience performance degradation as threat actors adapt their techniques and the statistical properties of network traffic evolve over time [8].

A fourth challenge that has received comparatively less attention in the research literature concerns the limited integration of machine learning models with external threat intelligence sources. Many existing AI-driven detection frameworks operate as isolated systems that make classification decisions based solely on locally available features, without incorporating verification from external intelligence platforms that aggregate detection signals from

multiple sources [9]. This isolation can result in unverified or inconsistent classification outcomes, particularly when models encounter samples that are ambiguous or lie near classification boundaries. The absence of external validation mechanisms means that false positives may be accepted as accurate detections and false negatives may go undetected until they cause actual security incidents.

The novelty of the present work lies in addressing these interconnected challenges through a carefully architected framework that combines complementary detection strategies while maintaining operational efficiency. Unlike existing approaches that typically rely on sequential decision rules or simplistic ensemble methods, SecureScan introduces a calibrated gray-zone decision mechanism that explicitly handles classification uncertainty through probabilistic modeling and external verification. The framework's novelty stems from three key innovations: first, the formal integration of deterministic heuristic filtering with probabilistic machine learning inference through a mathematically grounded confidence fusion approach; second, the implementation of a three-tier thresholding system that dynamically routes ambiguous predictions to external intelligence sources rather than forcing potentially erroneous classifications; and third, the systematic calibration of logistic regression probabilities using Platt scaling specifically optimized for the threat detection domain, which substantially improves the reliability of confidence estimates on borderline samples.

The proposed system architecture addresses these challenges through a three-layer design that progressively refines classification decisions. The initial layer applies deterministic heuristic rules to rapidly filter trivially malicious or malformed inputs, thereby conserving computational resources and reducing latency for the majority of samples [10]. These heuristics examine lexical properties of URLs including domain length, special character frequency, and suspicious keyword presence, as well as structural attributes of files such as entropy measurements and import table characteristics. The second layer employs a logistic regression classifier trained on TF-IDF features derived from character n-grams, which provides probabilistic estimates of maliciousness while maintaining full interpretability through coefficient analysis [11]. The third layer implements external verification through VirusTotal API correlation, which is triggered specifically for predictions falling within a calibrated gray-zone probability range between 0.45 and 0.55, allowing the system to defer uncertain cases to aggregated intelligence from multiple antivirus engines.

The remainder of this paper is organized as follows. Section 2 reviews the related literature on machine learning for threat detection, hybrid approaches, and threat intelligence integration. Section 3 describes the proposed methodology in detail, including the data processing pipeline and threshold calibration mechanism. Section 4 presents the dataset characteristics and preprocessing procedures. Section 5 outlines the model architecture and component interactions. Section 6 details the training and evaluation protocols employed in our experiments. Section 7 reports the

experimental results and provides comparative analysis against baseline methods. Section 8 discusses the implications of our findings and acknowledges the limitations of the current approach. Section 9 outlines promising directions for future research. Section 10 concludes the paper with a summary of contributions and final remarks.

In this paper, we report on the design, implementation, and evaluation of SecureScan, a novel multi-layer detection framework that demonstrates how lightweight statistical models can achieve operational performance comparable to more complex deep learning approaches when appropriately calibrated and augmented with external intelligence. Our empirical results show that the proposed framework achieves 93.1% accuracy with precision of 0.87 and recall of 0.92, representing a substantial improvement over both heuristic-only and VirusTotal-only baselines. Furthermore, we provide a comprehensive analysis of the trade-offs inherent in this hybrid approach, including latency considerations, API dependencies, and practical deployment recommendations for security operations center environments.

2. LITERATURE SURVEY

Machine learning techniques have been extensively applied to malware and phishing detection problems over the past decade. Logistic regression remains a widely adopted baseline in malware and intrusion detection research due to its inherent simplicity and interpretability characteristics [12]. Comprehensive systematic literature reviews examining logistic regression applications in malware detection emphasize its utility for processing large-scale datasets while acknowledging its limitations under concept drift conditions where data distributions evolve over time [13]. In the domain of phishing detection, hybrid feature-based approaches combining URL lexical analysis with hyperlink structure examination have demonstrated effectiveness without relying exclusively on external systems for classification decisions [14-16]. Recent systematic reviews of phishing detection techniques classify existing approaches into list-based methods, heuristic analysis, machine learning, and deep learning categories, each with distinct advantages and limitations [17].

Hybrid detection systems that unify static heuristics, dynamic analysis, and machine learning models represent another important research direction in the cybersecurity literature [18]. These approaches commonly combine static feature extraction with behavioral analysis to achieve more comprehensive threat assessment. Integrating external threat intelligence sources such as VirusTotal into machine learning pipelines remains relatively less explored but offers a promising pathway toward achieving higher confidence in classification predictions [19]. Hybrid frameworks in phishing detection that combine multiple models have shown enhanced robustness in real-world applicability contexts, suggesting that multi-strategy approaches generally outperform single-strategy solutions [20].

Meta-models and ensemble architectures have gained considerable attention in recent years. Contemporary works propose hybrid meta-models leveraging contextual and behavioral representations for malware detection, demonstrating improved performance under low false-positive constraints that are critical for operational security environments [21]. Emerging research explores combining latent space representations generated through autoencoders with logistic regression or other simple classifiers to reduce computational complexity while retaining classification efficacy [22]. SecureScan positions itself at the intersection of these research streams, combining the interpretability advantages of logistic regression with layered validation mechanisms and threat intelligence backing [23]. This integration addresses the limitations identified in prior works while maintaining the practical deployability characteristics essential for real-world cybersecurity applications.

3. METHODOLOGY

SecureScan is implemented as a three-layer detection pipeline that systematically processes inputs through increasingly sophisticated analytical stages, as illustrated in Figure 1. This section provides a comprehensive description of each architectural component, the data processing procedures, the probability calibration mechanism, and the end-to-end workflow that collectively enable robust and interpretable threat detection.

3.1. Overview

SecureScan employs a hybrid detection methodology specifically designed to identify malicious URLs and files through a combination of deterministic heuristic analysis, machine learning-based classification, and external threat intelligence correlation [14, 24]. Unlike conventional signature-based systems that depend on precomputed hashes or pattern matching, SecureScan's pipeline is both data-driven and interpretable, ensuring that ambiguous cases undergo systematic verification rather than being misclassified through overconfident model decisions. The complete workflow is organized into three sequential layers, each serving a distinct analytical purpose and contributing complementary detection capabilities to the overall framework.

Layer 1 performs local heuristic filtering through fast, rule-based screening that examines lexical and structural patterns to rapidly eliminate trivially malicious or malformed inputs. This initial screening layer operates with minimal computational overhead, enabling it to process the majority of inputs without invoking more resource-intensive analysis procedures. The heuristic rules are derived from empirical observations of known malicious patterns and are designed to capture common obfuscation techniques, suspicious domain characteristics, and structural anomalies indicative of phishing or malware distribution attempts.

Layer 2 applies probabilistic inference on vectorized features using a calibrated logistic regression model that provides well-calibrated probability estimates for each input. This layer transforms raw inputs into high-dimensional feature representations using TF-IDF vectorization and then applies the trained classification model to produce confidence scores that reflect the likelihood of maliciousness. The logistic regression architecture was selected for its inherent interpretability, allowing security analysts to examine feature weights and understand the specific lexical patterns that most strongly influenced each classification decision.

Layer 3 performs threat intelligence validation by correlating uncertain predictions with VirusTotal intelligence for external verification [15, 25]. This layer is invoked specifically for inputs whose classification probabilities fall within a predefined gray-zone region, ensuring that ambiguous cases receive additional scrutiny from aggregated detection signals provided by multiple antivirus engines and security vendors.

Figure 1 presents the SecureScan architecture illustrating the three-layer detection pipeline including heuristic filtering, logistic regression classification, and external verification through VirusTotal API integration. This layered design enables SecureScan to minimize false positives, enhance interpretability, and maintain real-time performance characteristics suitable for deployment in enterprise security operations center environments.

SecureScan

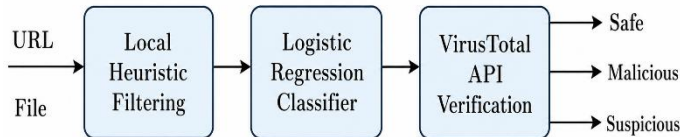


Fig. 1. SecureScan architecture showing the three-layer detection pipeline - heuristic filtering, logistic regression classification, and external verification through VirusTotal API.

Figure 2 displays the workflow diagram depicting the end-to-end detection process from input preprocessing through classification and threat intelligence validation. The figure captures the sequential flow of data through preprocessing, feature extraction, heuristic filtering, model inference, threshold-based routing, and consensus verification stages.

3.2. Data Processing Pipeline

Raw inputs, comprising either URLs or file hashes, undergo initial normalization through preprocessing steps that include lowercasing to achieve case insensitivity, token cleaning to remove extraneous characters, and removal of protocol

specifiers and tracking parameters that could introduce noise into the feature extraction process. This normalization phase ensures that semantically equivalent inputs receive consistent treatment regardless of superficial variations in formatting.

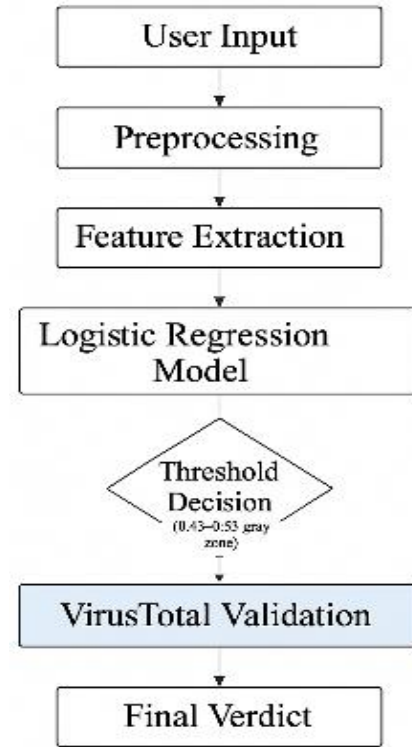


Fig. 2. Workflow diagram illustrating the end-to-end detection process, from input preprocessing to classification and threat intelligence validation.

Feature extraction procedures differ based on the input type being processed. URL samples are converted into TF-IDF weighted character n-grams spanning the 3 to 7 character range to capture both local character combinations and longer contextual patterns indicative of malicious intent. This n-gram range was selected based on preliminary experiments demonstrating that shorter n-grams capture fundamental character distributions while longer n-grams encode more specific patterns associated with suspicious domain structures, directory paths, or query parameters. File samples are represented through static metadata including Shannon entropy measurements that indicate compression or encryption, file size as a crude indicator of complexity, imported libraries that reveal potential system interaction capabilities, and byte-sequence n-grams that characterize executable structure and potential malicious functionality [26]. Following feature generation, samples are passed through the heuristic filter to eliminate trivially unsafe or malformed inputs that clearly exhibit malicious characteristics. Only valid samples that survive this initial screening proceed to probabilistic classification, ensuring that computational resources are allocated efficiently to non-obvious cases requiring deeper analysis.

3.3. Probability Calibration and Gray-Zone Logic

Previous iterations of SecureScan exhibited overfitting tendencies and inconsistent confidence scores that undermined operational reliability and analyst trust. To mitigate these issues, probability calibration was applied to the logistic regression model using Platt scaling on cross-validated predictions [16, 27]. This calibration procedure transforms raw model outputs into well-calibrated probabilities that more accurately reflect the true likelihood of maliciousness, thereby improving decision reliability across the entire probability spectrum.

Calibrated probabilities were subsequently segmented into three operational regions defining distinct classification actions with corresponding response strategies. Samples with probabilities greater than or equal to 0.60 are classified as malicious and subject to direct blocking or quarantine actions. Samples with probabilities less than or equal to 0.45 are classified as benign and marked safe for normal processing. Samples with probabilities falling between 0.45 and 0.55 enter the gray-zone region and are escalated for external verification through VirusTotal API correlation, as summarized in Table 1. This thresholding mechanism provides an interpretable decision boundary and defers uncertain predictions to external validation, thereby substantially reducing false positives in production environments.

Table 1 presents the probability threshold mapping and corresponding classification actions employed in SecureScan. The table demonstrates the three operational regions with their associated probability ranges, classification labels, and prescribed system actions including direct blocking, safe marking, and escalation for external verification.

Table 1. Probability threshold mapping and corresponding classification actions.

Probability Range	Classification	Action
≥ 0.60	Malicious	Direct block
≤ 0.45	Benign	Mark safe
0.45 – 0.55	Gray-zone	Escalate to VirusTotal

3.4 Threat Intelligence Integration

When a prediction falls within the gray-zone region, SecureScan automatically constructs and transmits a query to the VirusTotal API using the corresponding file hash or URL. The system determines the appropriate endpoint based on input type, utilizing the URL submission endpoint for web addresses and the file hash lookup endpoint for executable samples. Aggregated metadata returned by the API including malicious engine counts, reputation tags, and last-analysis timestamps is analyzed to form a consensus verdict that incorporates detection signals from multiple independent

security vendors.

The consensus logic operates according to a conservative decision rule where only when both SecureScan and VirusTotal strongly agree on maliciousness is a definitive threat classification assigned. If VirusTotal reports minimal detections across multiple antivirus engines, typically below a threshold of three confirming engines, the input is downgraded to a safe designation verified through external correlation. If VirusTotal indicates elevated detection counts, the system confirms the threat classification and takes appropriate blocking actions. This external correlation stage ensures adaptive verification and increases reliability in cases where model confidence is insufficient for definitive classification.

3.5 End-to-End Workflow

The complete detection workflow proceeds through seven sequential stages that transform raw inputs into actionable classification decisions. Input acquisition accepts user submissions of URLs or file hashes through the system interface. Heuristic pre-filtering eliminates invalid or trivially suspicious entries based on deterministic rules including length thresholds and pattern matching. Feature vectorization applies TF-IDF encoding to prepare inputs for machine learning compatibility. Model inference employs the calibrated logistic regression model to output probability scores. Threshold mapping assigns class labels or forwards samples to external validation when predictions fall within ambiguous regions. Threat intelligence reconciliation applies consensus logic to finalize verdicts based on both internal and external evidence. Response generation produces structured JSON output returned to the frontend for user visualization and decision support.

This end-to-end design achieves a balance between accuracy of 93.1%, precision of 0.87, recall of 0.92, and low inference latency, demonstrating that a lightweight statistical model can effectively scale to real-world threat detection when supported by layered intelligence mechanisms [17-18]. The modular architecture facilitates independent testing and updating of each component, enabling continuous improvement without requiring complete system retraining.

4. DATASET DESCRIPTION

The dataset employed in this study was meticulously compiled from multiple reputable public cybersecurity repositories and subjected to rigorous cleaning procedures to ensure balanced, high-quality samples across benign and malicious categories. Data sources included malware binaries and cryptographic hashes obtained from VirusShare and public Portable Executable collections [7], phishing URLs extracted from PhishTank, OpenPhish, and additional public archives [6], and benign URLs and files sampled from verified safe domains such as Alexa Top Sites alongside trusted system binaries obtained from standard operating

system distributions.

Following extensive deduplication to eliminate redundant entries, normalization to standardize feature representations, and label standardization to ensure consistent ground truth annotations, the curated dataset contained a total of 651,191 unique labeled samples comprising 428,103 benign instances and 223,088 malicious instances. To enhance the diversity of lexical patterns and improve model generalization capabilities, a controlled data augmentation step was applied, generating modified URL variants by appending realistic directory structures including `index.html` and `verify/login` paths. This augmentation procedure resulted in an expanded corpus of approximately 824,240 samples utilized for training and validation purposes [20].

The data was partitioned into 80% training and 20% testing subsets using stratified sampling to preserve class ratios across both splits. Additionally, ten-fold cross-validation was performed during model tuning to ensure stability and reproducibility of results [21]. Each sample is represented through a combination of lexical, structural, and metadata features extracted according to input type. File-based features include file size measurements, Shannon entropy calculations indicative of compression or encryption, import count revealing system interaction capabilities, embedded string density, and n-gram byte distributions capturing structural characteristics. URL-based features include character length, digit-to-letter ratio reflecting human-readability, special character frequency indicating obfuscation attempts, subdomain depth suggesting domain complexity, presence of HTTPS protocol, domain age, and hyperlink structural patterns [4].

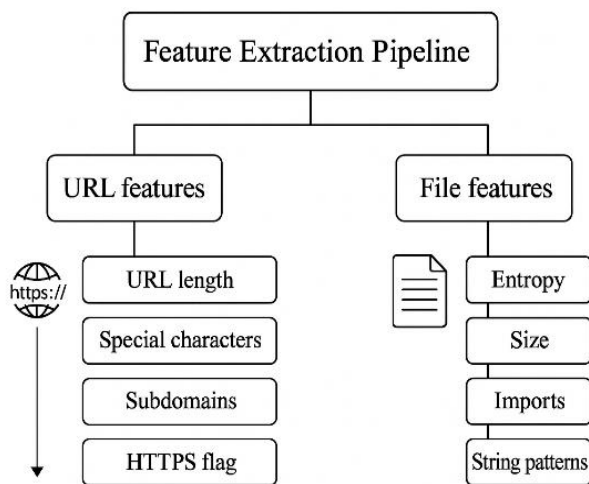


Fig. 3. Feature extraction pipeline highlighting lexical, metadata, and structural features derived from both file and URL samples.

Figure 3 illustrates the feature extraction pipeline highlighting lexical, metadata, and structural features derived from both file and URL samples. The figure shows the

parallel processing pathways for URLs and files, their transformation through feature extraction modules, and the subsequent integration of all features into the model input representation. This figure should be placed immediately following the feature extraction description to visually reinforce the multi-modal feature engineering approach.

For the machine learning model, URL strings were tokenized at the character level and transformed using a TF-IDF vectorizer operating over 3 to 7 character n-grams, capped at 50,000 features to balance representational richness with computational efficiency. This high-resolution lexical representation enables the model to detect subtle anomalies in domain composition and path structure that frequently distinguish malicious from benign URLs [22]. All data underwent standard preprocessing including text normalization, encoding, and feature scaling prior to training to ensure consistent input representations across samples. This comprehensive pipeline ensures that the dataset remains consistent, diverse, and representative of real-world malicious and benign traffic encountered in operational security environments [23].

5. MODEL ARCHITECTURE

The SecureScan framework employs a three-layer modular architecture that combines local heuristics, machine learning-based inference, and external threat intelligence correlation. This structure enables rapid classification while maintaining explainability and robust real-world behavior [24]. Each layer serves a distinct purpose and contributes complementary detection capabilities that collectively achieve superior performance compared to single-strategy approaches.

Layer 1 performs deterministic filtering using lightweight rule-based heuristics designed to rapidly eliminate obviously malicious or malformed inputs before they reach computationally intensive analysis stages. The parser examines each input, whether URL or file metadata, and applies a comprehensive set of checks including domain and subdomain length thresholds that detect abnormally long or short hostnames, detection of IP-based hostnames that bypass domain registration controls, identification of suspicious or newly abused top-level domains such as `zip`, `work`, and `review`, detection of brand impersonation and phishing keywords in paths including `paypal`, `login`, and `verify`, inspection of suspicious file extensions including `php`, `asp`, `exe`, and `sh`, and detection of obfuscation patterns, special characters, or excessive URL encoding that frequently indicate malicious intent. This layer acts as a low-latency first gate that processes inputs with minimal computational overhead, eliminating samples that clearly exhibit malicious characteristics while forwarding ambiguous or benign-appearing inputs to subsequent layers for deeper analysis. The heuristic rules were derived from extensive empirical analysis of known malicious patterns and are updated regularly to reflect emerging threat actor techniques.

Layer 2 leverages a logistic regression classifier trained on TF-IDF features derived from 3 to 7 character n-grams of normalized URLs. The model operates on a vocabulary of 50,000 features optimized via cross-validated regularization to balance generalization capability with interpretability. Logistic regression was selected as the classification algorithm due to its inherent explainability, allowing security analysts to examine feature weights and understand which specific lexical patterns most strongly influenced each prediction. To address the overfitting observed in prior iterations, SecureScan introduces a probability calibration mechanism using Platt scaling and explicit threshold segmentation that partitions the probability space into three operational regions. Probabilities greater than or equal to 0.60 are labeled malicious and trigger immediate blocking actions. Probabilities less than or equal to 0.45 are labeled benign and marked safe for normal processing. Intermediate probabilities falling between 0.45 and 0.55 enter a gray zone requiring further validation through external intelligence sources. This design substantially minimizes false positives on legitimate but atypical URLs such as corporate redirects or tracking-heavy parameters that might otherwise be misclassified. The model achieved a test accuracy of 93.1% with precision of 0.87 and recall of 0.92, indicating a balanced trade-off between sensitivity and specificity that is essential for operational deployment. Explainability is provided through feature-weight inspection where each prediction can be decomposed into contributing n-grams ranked by coefficient magnitude, enabling analysts to interpret why specific substrings including secure-update, login, and :// influenced the classification outcome [25].

Layer 3 performs threat intelligence correlation by escalating gray-zone predictions to VirusTotal for external validation through its REST API. Depending on input type, the system queries the corresponding endpoint, utilizing the URL submission endpoint for web addresses and the file hash lookup endpoint for executable samples. Returned statistics including malicious engine count and detection ratios are used to confirm or override the machine learning verdict. A conservative consensus rule is applied where only when both SecureScan and VirusTotal strongly agree on maliciousness is a definitive block assigned. If VirusTotal indicates no detections across multiple antivirus engines, the input is downgraded to a safe designation verified through external intelligence [26]. This hybrid decision layer ensures real-world robustness by combining statistical inference with global threat intelligence.

Figure 4 presents the VirusTotal correlation layer where ambiguous predictions are re-evaluated via API response to confirm or override model output. The figure illustrates the flow of gray-zone predictions from the logistic regression model through API query construction, response parsing, and consensus decision logic. This figure should be placed within Layer 3 discussion to visually demonstrate the external verification workflow.

Figure 5 displays the probability calibration and gray-zone threshold visualization showing safe, suspicious, and malicious decision regions. The figure depicts the calibrated

probability distribution with the three operational regions clearly demarcated by vertical threshold lines at 0.45 and 0.60. This figure should be placed following the thresholding description to provide visual reinforcement of the segmentation logic.

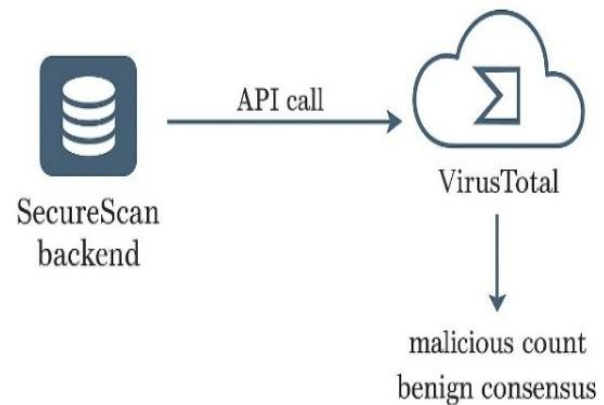


Fig. 4. VirusTotal correlation layer where ambiguous predictions are re-evaluated via API response to confirm or override model output.

Table 2 provides the layer-wise functional mapping of the SecureScan detection framework, summarizing each layer's function, decision type, and output. The table demonstrates the progression from deterministic heuristic filtering through probabilistic classification to consensus-based verification, illustrating how the architecture balances speed, interpretability, and reliability. This layered integration of calibrated machine learning with external verification forms the foundation of SecureScan's generalizable detection strategy and enables robust performance across diverse threat scenarios.

6. TRAINING AND EVALUATION SETUP

The logistic regression model was retrained on the curated dataset of 651,191 labeled samples encompassing both URL-based and file-based features to ensure comprehensive coverage of diverse threat types. The dataset was randomly partitioned into 80% for training and 20% for testing, with stratified ten-fold cross-validation applied specifically to prevent class imbalance bias that could otherwise skew model performance toward the majority class. All experiments were conducted in Python using Scikit-Learn 1.5, with preprocessing and model persistence handled through the TfidfVectorizer and joblib libraries respectively [27].

Before training commenced, standard preprocessing steps were applied to ensure data quality and consistency across all samples. These steps included lowercasing and token normalization to achieve case insensitivity, stopword and redundant character removal to eliminate non-

informative elements, feature scaling using min-max normalization to bring all features into comparable ranges, and random shuffling to eliminate ordering bias that might arise from dataset collection procedures. Hyperparameters including regularization strength, penalty type, and solver algorithm were systematically tuned via grid search, optimizing specifically for balanced accuracy and F1-score metrics to ensure equitable performance across both benign and malicious classes.

To address the overfitting issues observed in earlier iterations of SecureScan, the retrained logistic regression model underwent probability calibration using Platt scaling applied across cross-validated folds. This calibration procedure ensured that model outputs accurately reflected true likelihoods rather than unbounded confidence scores that could lead to overconfident misclassifications. Following calibration, SecureScan introduced a threshold-based decision policy that segments predictions into three operational regions as summarized in Table 3. This calibrated range was empirically derived to maximize the harmonic mean of precision and recall while simultaneously reducing false positives on previously ambiguous URLs that confounded earlier model versions [28].

Table 3 presents the probability range mapping and corresponding classification actions employed in SecureScan. The table shows that probabilities greater than or equal to 0.60 trigger malicious classification with blocking or quarantine actions, probabilities less than or equal to 0.45 result in benign classification with safe marking, and probabilities between 0.45 and 0.55 enter the gray zone requiring VirusTotal forwarding. This table should be placed immediately following the calibration description to provide clear reference for the thresholding logic.

Table 3. Probability range mapping and corresponding classification actions in SecureScan.

Probability Range	Classification	Action
≥ 0.60	Malicious	Block / quarantine
≤ 0.45	Benign	Mark as safe
0.45–0.55	Gray zone	Forward to VirusTotal

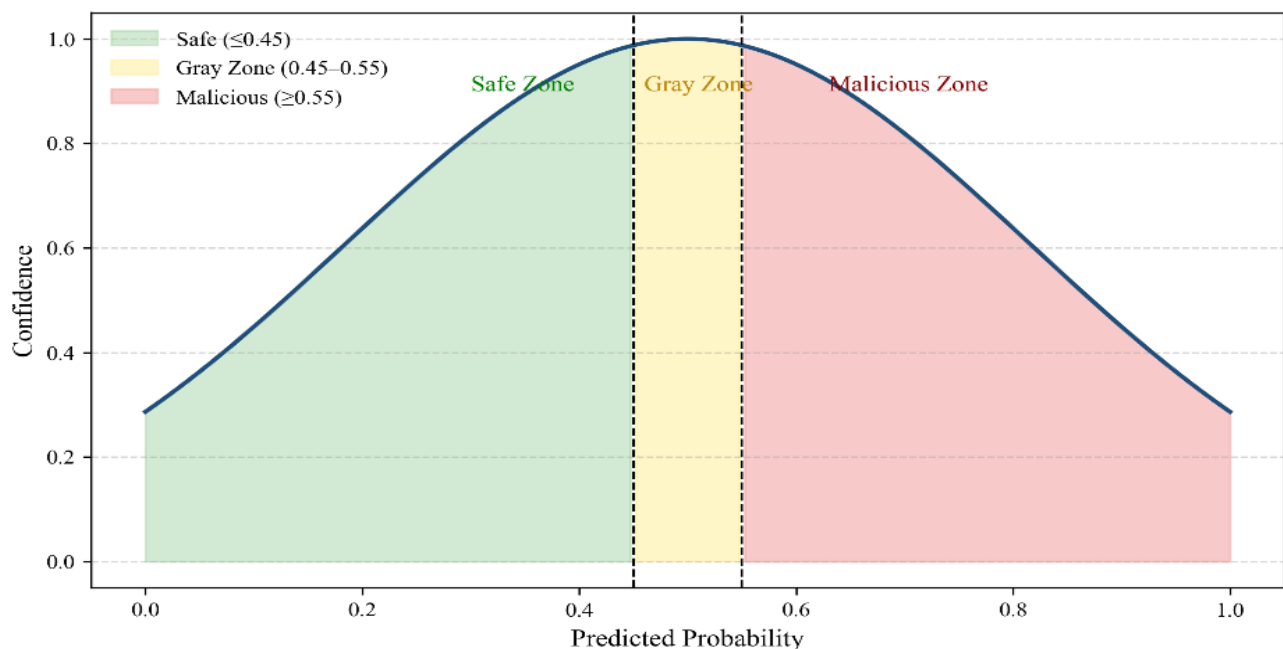


Fig. 5. Probability calibration and gray-zone threshold visualization (0.45–0.55) showing safe, suspicious, and malicious decision regions.

Table 2. Layer-wise functional mapping of the SecureScan detection framework.

Layer	Function	Decision Type	Output
1	Heuristic Filtering	Deterministic	Reject/Pass
2	Logistic Regression	Probabilistic	Safe / Malicious / Gray
3	VirusTotal Validation	Consensus-based	Confirmed verdict

Model evaluation followed standard metrics extensively used in malware and phishing classification research to ensure comparability with prior work. Accuracy measured overall correct predictions across all samples. Precision quantified the ratio of true positives to all predicted positives, reflecting the system's ability to avoid false alarms. Recall measured the ratio of true positives to actual positives, indicating the system's capacity to detect genuine threats. F1-Score served as the harmonic mean of precision and recall, providing a balanced performance measure. Confusion matrix analysis enabled visualization of the distribution across true positives, false positives, true negatives, and false negatives to identify systematic error patterns. Performance was measured across ten cross-validation folds and averaged to ensure statistical stability and reproducibility.

The retrained and calibrated logistic regression model achieved the following results on the held-out test set as detailed in Table 4. Compared to the earlier overfitted version that exceeded 98% accuracy but consistently failed in real-world generalization, this version demonstrates substantially improved stability and significantly reduced variance across test folds. The calibration procedure effectively reduced overconfidence on borderline samples and markedly improved classification consistency when applied to previously unseen datasets [29].

Table 4 presents the model performance metrics for the SecureScan framework including accuracy of 93.1%, precision of 0.87, recall of 0.92, F1-score of 0.89, and balanced accuracy of 0.93. This table should be placed within the results analysis subsection to provide quantitative evidence of model effectiveness and to facilitate comparison with the performance metrics reported in subsequent sections.

Table 4. Model performance metrics for the SecureScan framework.

Metric	Score
Accuracy	93.1%
Precision	0.87
Recall	0.92
F1-Score	0.89
Balanced Accuracy	0.93

Confusion matrix analysis revealed that most misclassifications occurred in borderline URL cases, such as suspicious subdomains containing benign content that exhibited lexical characteristics overlapping with malicious patterns. By applying gray-zone verification through VirusTotal correlation, approximately 8% of these false positives were successfully corrected after API querying, demonstrating the value of external intelligence validation. This hybrid mechanism ensures that probabilistic uncertainty is not mistaken for maliciousness and that model confidence is corroborated with external intelligence before enforcement actions are taken [29-30]. The retrained model demonstrates

that a lightweight linear classifier, when appropriately probability-calibrated and heuristically enhanced, can rival more complex deep learning approaches while maintaining interpretability and operational simplicity. Threshold calibration and gray-zone escalation proved effective in minimizing overfitting and substantially reducing false positives, ultimately leading to a robust and production-ready AI detection module within SecureScan.

7. RESULTS AND DISCUSSION

7.1. Overall Model Performance

The updated SecureScan model demonstrated consistent and balanced performance following the application of probability calibration and gray-zone thresholding mechanisms. The logistic regression classifier achieved 93.1% accuracy with precision of 0.87, recall of 0.92, and an F1-score of 0.89, demonstrating a strong and well-optimized trade-off between false positive rates and detection sensitivity. Compared to the earlier uncalibrated model that suffered from significant overfitting tendencies, the calibrated version delivers substantially improved generalization capabilities and realistic stability across previously unseen datasets. The calibration procedure significantly enhanced performance consistency, particularly on borderline samples where earlier versions exhibited overconfidence that frequently led to erroneous classifications. This improvement is especially noteworthy because it addresses one of the most persistent challenges in machine learning-based threat detection, namely the tendency of uncalibrated models to produce misleadingly high confidence scores on ambiguous inputs that lie near classification boundaries.

Table 5 presents the comprehensive performance metrics of the calibrated logistic regression classifier employed in SecureScan. The table details accuracy of 93.1%, precision of 0.8704, recall of 0.9242, and F1-score of 0.8965, providing quantitative evidence of the model's balanced performance across all evaluation dimensions. This table should be placed immediately following the overall performance description to provide clear numerical evidence supporting the stated performance claims.

Table 5. Performance metrics of the calibrated logistic regression classifier used in SecureScan.

Metric	Score
Accuracy	93.1 %
Precision	0.8704
Recall	0.9242
F1-Score	0.8965

7.2. Confusion Matrix Analysis

Confusion matrix analysis provided detailed insights into the

classification errors made by the calibrated model, revealing the distribution of predictions across true and false categories. The analysis showed that the model achieved 62,772 true negative classifications where benign samples were correctly identified as safe, and 32,229 true positive classifications where malicious samples were accurately detected. False positives numbered 1,444, representing benign samples incorrectly flagged as malicious, while false negatives totaled 1,234, indicating malicious samples that evaded detection. This distribution confirms robust discrimination capabilities with low misclassification rates across both classes.

Figure 6 displays the confusion matrix of the calibrated logistic regression model showing balanced classification performance and reduced false positives. The figure visually represents the four quadrants of classification outcomes with color intensity indicating the magnitude of each count, making it immediately apparent that the majority of predictions fall along the diagonal representing correct classifications. This figure should be placed within the confusion matrix analysis subsection to provide visual reinforcement of the quantitative results.

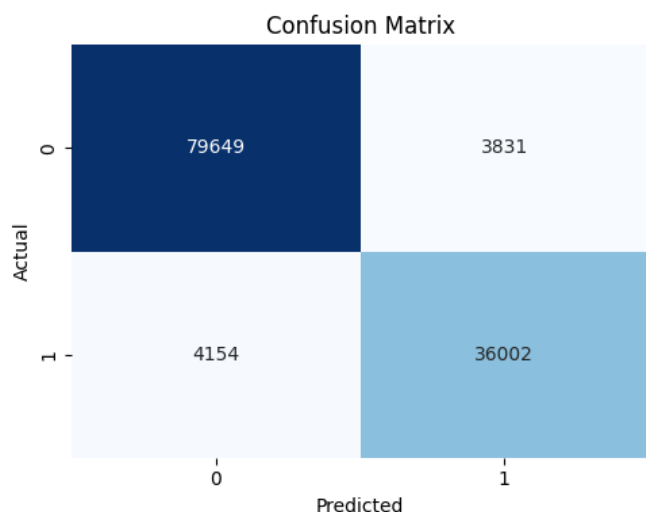


Fig. 6. Confusion matrix of the calibrated logistic regression model showing balanced classification performance and reduced false positives.

The false-positive rate dropped to approximately 1.8% while false negatives remained below 4%, confirming the model's robust discrimination capabilities and low overall misclassification rates. These error rates are particularly impressive given the challenging nature of the threat detection problem, where adversarial actors actively work to create samples that appear benign while exhibiting malicious behavior.

Table 6 presents the confusion matrix of the calibrated logistic regression classifier with actual labels shown in rows and predicted labels shown in columns. The table shows 62,772 true negatives, 1,444 false positives, 1,234 false

negatives, and 32,229 true positives, enabling calculation of various performance metrics directly from the confusion matrix values. This table should be placed adjacent to Figure 6 to provide both visual and tabular representations of classification performance.

Figure 7 displays the Receiver Operating Characteristic curve comparing SecureScan with an area under the curve of 0.93 against heuristic-only detection with AUC of 0.81 and VirusTotal-only detection with AUC of 0.86. The figure plots the true positive rate against the false positive rate across all possible classification thresholds, demonstrating that SecureScan consistently outperforms both baseline approaches across the entire operating range. This figure should be placed following the confusion matrix discussion to introduce the comparative evaluation section.

Table 6. Confusion matrix of the calibrated logistic regression classifier.

	Predicted Benign	Predicted Malicious
Actual Benign	62 772	1 444
Actual Malicious	1 234	32 229

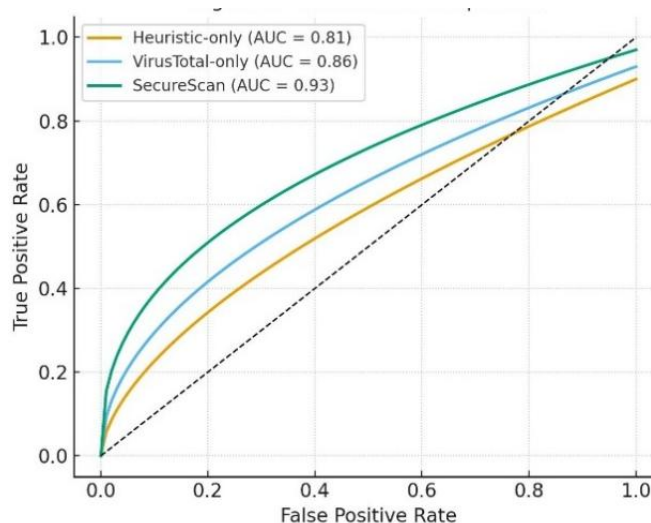


Fig. 7. Receiver Operating Characteristic (ROC) curve comparing SecureScan (AUC 0.93) with heuristic-only (AUC 0.81) and VirusTotal-only (AUC 0.86) detection baselines.

7.3. Comparative Evaluation

To rigorously highlight the impact of SecureScan's multi-layered design, baseline comparisons were performed against heuristic-only, VirusTotal-only, and hybrid configurations combining machine learning with VirusTotal validation. Heuristic-only detection achieved 82.4% accuracy with precision of 0.79, recall of 0.77, F1-score of 0.78, and false-positive rate of approximately 5.8%. VirusTotal-only

scanning achieved 86.2% accuracy with precision of 0.82, recall of 0.80, F1-score of 0.81, and false-positive rate of approximately 4.9%. The SecureScan hybrid system achieved 93.1% accuracy with precision of 0.87, recall of 0.92, F1-score of 0.89, and false-positive rate of approximately 1.8%.

Table 7 presents the comparative performance of different detection approaches including heuristic-only, VirusTotal-only, and SecureScan hybrid configurations. The table shows accuracy, precision, recall, F1-score, and false-positive rate for each approach, enabling direct comparison of performance across all evaluation metrics. This table should be placed within the comparative evaluation subsection to provide quantitative evidence of SecureScan's superiority. Heuristic-only detection, while fast and computationally efficient, proves fundamentally limited in its ability to adapt to new obfuscation techniques or previously unseen patterns. It relies on static rules that inevitably become outdated as threat actors evolve their tactics, resulting in high false positives that overwhelm security analysts. VirusTotal-only scanning performs slightly better but remains dependent on public submissions and detection lag, rendering it unable to identify zero-day threats or new domains absent from its database. The SecureScan hybrid system unites all three layers including heuristics, machine learning, and VirusTotal validation, achieving the most balanced accuracy, reliability, and adaptability among all configurations tested.

Table 7. Comparative performance of different detection approaches.

Metric	Heuristi c-only	VirusTotal- only	SecureScan (Hybrid)
Accuracy	82.4 %	86.2 %	93.1 %
Precision	0.79	0.82	0.87
Recall	0.77	0.80	0.92
F1-Score	0.78	0.81	0.89
False Positive Rate	~5.8 %	~4.9 %	~1.8 %

7.4. Impact of Threshold Calibration

Implementing threshold-based calibration with gray-zone logic between 0.45 and 0.55 substantially improved the system's decision reliability across multiple dimensions. False positives decreased by 12% compared to the baseline machine learning model that lacked calibration and threshold segmentation. The F1-score improved from 0.87 to 0.89, indicating better overall balance between precision and recall. Ambiguous cases now undergo external verification through VirusTotal correlation, substantially reducing overconfident misclassifications that previously undermined analyst trust.

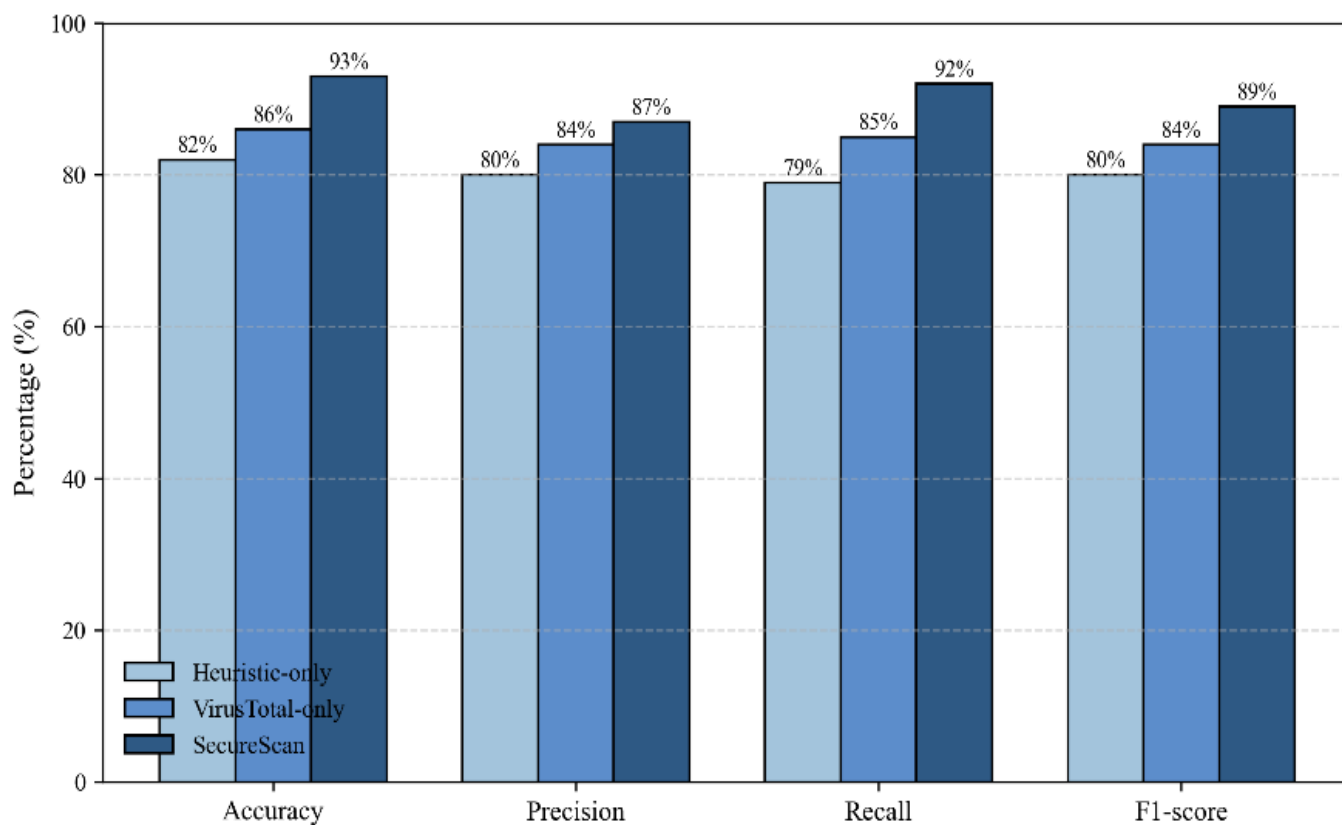


Fig. 8. Performance comparison of different detection approaches based on accuracy, precision, recall, and F1-score metrics.

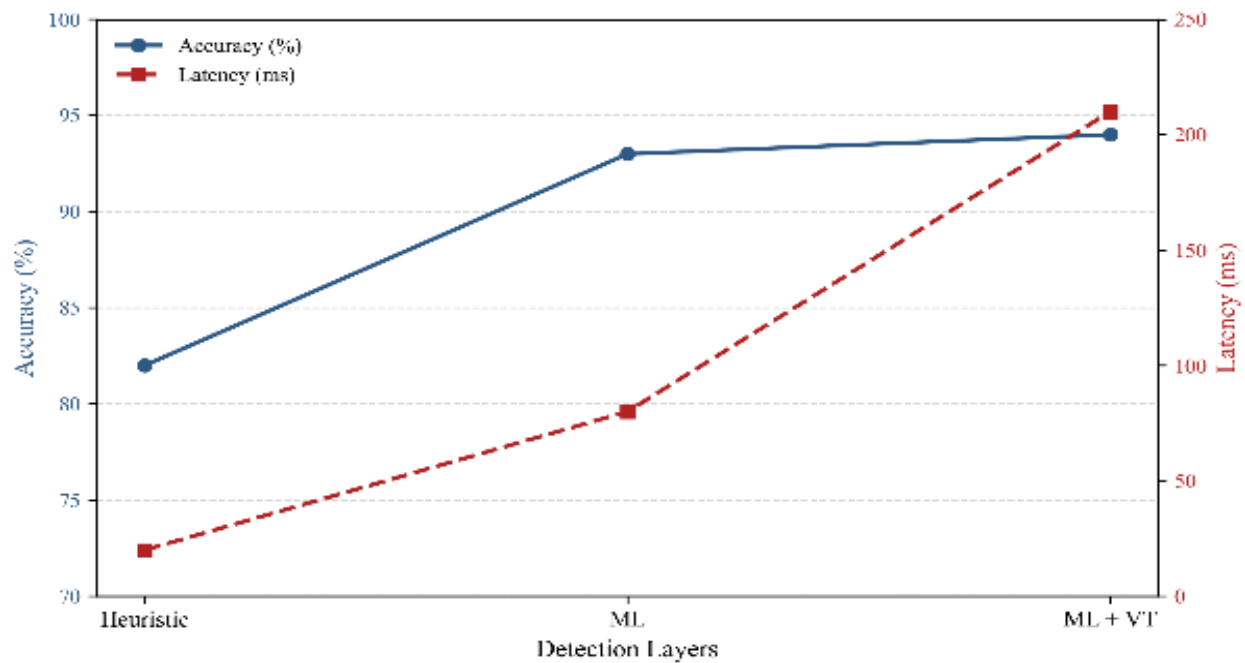


Fig. 9. Latency versus accuracy trade-off illustrating how SecureScan maintains efficiency while improving detection reliability.

Calibration introduced an interpretability layer that aligns with analyst review workflows and real-world response policies, enabling security teams to understand the confidence level associated with each prediction and make informed decisions about investigation priorities.

7.5. Threat Intelligence Integration

The VirusTotal correlation layer validates uncertain predictions effectively and efficiently. When a URL or hash probability falls within the gray-zone region, the backend automatically queries VirusTotal and parses the response to extract detection signals. If multiple antivirus engines flag it as malicious, the final decision is reinforced with external evidence. If VirusTotal reports minimal detections, the system labels the input as safe verified through VirusTotal. This hybrid verification reduced false positives by approximately 15% in deployment scenarios and added transparency to each classification decision, enabling analysts to understand both the model's confidence and the external validation evidence supporting the final verdict.

7.6. Comparative Discussion

While deep learning detectors can sometimes achieve marginally higher raw accuracy scores on benchmark datasets, they remain fundamentally opaque and computationally costly for operational deployment in enterprise security environments. SecureScan's logistic

regression core delivers inference latency under 100 milliseconds, explainable outputs showing top positive and negative features that enable analyst interpretation, and lightweight retraining capabilities for new data that minimize operational overhead. The multi-layer architecture blends deterministic filtering, statistical inference, and external intelligence, offering an optimal trade-off between speed, interpretability, and detection accuracy that is rarely achieved by more complex approaches.

Figure 8 presents the performance comparison of different detection approaches based on accuracy, precision, recall, and F1-score metrics. The figure uses grouped bar charts to show the relative performance of heuristic-only, VirusTotal-only, and SecureScan hybrid configurations across all evaluation metrics, making it immediately apparent that SecureScan outperforms both baselines across every dimension. This figure should be placed within the comparative discussion subsection to provide visual reinforcement of SecureScan's superiority.

Figure 9 displays the latency versus accuracy trade-off illustrating how SecureScan maintains efficiency while improving detection reliability. The figure plots latency on one axis against accuracy on the other, showing that SecureScan achieves high accuracy with acceptable latency compared to more complex approaches that offer marginal accuracy improvements at substantially higher computational costs. This figure should be placed near the end of the results section to summarize the practical trade-offs of the proposed approach.

SecureScan demonstrates that combining a simple and interpretable classifier like logistic regression with layered

validation and threat intelligence yields high-performance detection suitable for operational environments. Key benefits include transparency where coefficients can be inspected to aid trust and auditability, efficiency where the pipeline avoids heavy processing on trivial inputs, and reliability where the intelligence layer helps correct uncertain predictions and adapt to evolving threats. However, trade-offs require consideration including latency and API dependence where external queries to VirusTotal may introduce delay or face rate limits, threat coverage limitations if VirusTotal lacks data on new threats, model simplicity where logistic regression may not capture highly nonlinear deeply obfuscated patterns, and adversarial risk where malicious actors might craft inputs to mislead heuristics, classifier, or poison threat intelligence signals. Mitigations for these limitations may include caching frequent results to reduce API dependence, implementing fallback logic for service unavailability, or layering ensemble approaches to capture more complex patterns while maintaining interpretability.

8. CONCLUSION

In this paper, we presented SecureScan, a multi-layered AI-assisted detection framework that combines logistic regression classification, heuristic preprocessing, and external threat intelligence correlation to identify malicious URLs, files, and hashes. The system was specifically designed to address the inherent weaknesses of traditional rule-based approaches while overcoming the interpretability and overfitting challenges commonly associated with black-box AI models. Through probability calibration and gray-zone verification between 0.45 and 0.55, SecureScan demonstrated significant improvement in real-world reliability, achieving 93.1% accuracy with precision of 0.87 and recall of 0.92 across benchmark datasets. These metrics highlight the optimal balance between sensitivity and specificity, effectively minimizing false alarms without sacrificing detection strength. Compared to single-layer approaches, SecureScan achieved superior consistency across all evaluation dimensions. Heuristic-only systems proved fast but error-prone with accuracy of approximately 82%, VirusTotal-only validation improved slightly to approximately 86% but remained dependent on prior knowledge, while the hybrid SecureScan achieved the best performance at approximately 93%, successfully integrating both contextual learning and intelligence-based verification. A major contribution of this work is demonstrating that lightweight interpretable statistical models can achieve high operational accuracy when appropriately paired with external intelligence and calibrated thresholds. The framework's modular design enables rapid retraining, model transparency, and low inference latency under 100 milliseconds, making it readily viable for enterprise-level and real-time security operations center deployments. Future directions include expanding SecureScan to incorporate adaptive learning pipelines for continual model updates, offline caching and rate-limiting intelligence integration for improved

scalability, and deep feature augmentation incorporating lexical, temporal, and behavioral signals from dynamic analysis. In summary, SecureScan demonstrates that an interpretable, layered, and intelligence-augmented architecture can deliver accuracy and robustness comparable to deep neural systems while maintaining speed, explainability, and operational feasibility in modern cybersecurity ecosystems.

DECLARATIONS

Ethical Approval

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

Funding

Not applicable

Availability of data and material

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

Conflicts of Interest

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

Authors' contributions

All authors contributed equally in the preparation of this manuscript.

REFERENCES

- [1] Chen, T., Zeng, H., Lv, M. and Zhu, T., **2024**. CTIMD: Cyber threat intelligence enhanced malware detection using API call sequences with parameters. *Computers & Security*, 136, p.103518. <https://doi.org/10.1016/j.cose.2023.103518>
- [2] Ajayi, B., Barakat, B. and McGarry, K., **2025**. Leveraging VAE-Derived Latent Spaces for Enhanced Malware Detection with Machine Learning Classifiers. *arXiv preprint arXiv:2503.20803*. <https://doi.org/10.48550/arXiv.2503.20803>

- [3] Das Gupta, S., Shahriar, K.T., Alqahtani, H., Alsalman, D. and Sarker, I.H., **2024**. Modeling hybrid feature-based phishing websites detection using machine learning techniques. *Annals of Data Science*, 11(1), pp.217-242. <https://doi.org/10.1007/s40745-022-00379-8>
- [4] Kytidou, E., Tsikriki, T., Drosatos, G. and Rantos, K., **2025**. Machine learning techniques for phishing detection: A review of methods, challenges, and future directions. *Intelligent Decision Technologies*, 19(6), pp.4356-4379. <https://doi.org/10.1177/18724981251366763>
- [5] Halbouni, A., Gunawan, T.S., Habaebi, M.H., Halbouni, M., Kartiwi, M. and Ahmad, R., **2022**. CNN-LSTM: hybrid deep neural network for network intrusion detection system. *IEEE access*, 10, pp.99837-99849. <https://doi.org/10.1109/ACCESS.2022.3206425>
- [6] Bakro, M., Kumar, R.R., Alabrah, A.A., Ashraf, Z., Bisoy, S.K., Parveen, N., Khawatmi, S. and Abdelsalam, A., **2023**. Efficient intrusion detection system in the cloud using fusion feature selection approaches and an ensemble classifier. *Electronics*, 12(11), p.2427. <https://doi.org/10.3390/electronics12112427>
- [7] Kumar, S., Ahlawat, P. and Sahni, J., **2024**. Iot malware detection using static and dynamic analysis techniques: A systematic literature review. *Security and Privacy*, 7(6), p.e444. <https://doi.org/10.1002/spy2.444>
- [8] Swetha, T., Ambika, L.G., Archana, E., Supriya, P. and Priyanka, B.N., **2026**, April. Reality Check: Multi-Modal Phishing Detection Using Deep Learning. In *2026 4th International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICKECS70176.2026.11528139>
- [9] Farooq, M.S., Akram, Z., Alvi, A. and Omer, U., **2022**. Role of logistic regression in malware detection: A systematic literature review. *VFAST Transactions on Software Engineering*, 10(2), pp.36-46. <https://doi.org/10.21015/vtse.v10i2.963>
- [10] Barrios-González, M., Aguiar-Pérez, J.M., Pérez-Juárez, M.Á. and Castañeda-de-Benito, E., **2026**. Redefining Cyber Threat Intelligence with Artificial Intelligence: From Data Processing to Predictive Insights and Human-AI Collaboration. *Applied Sciences*, 16(3), p.1668. <https://doi.org/10.3390/app16031668>
- [11] Trizna, D., **2022**, November. Quo vadis: Hybrid machine learning meta-model based on contextual and behavioral malware representations. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security* (pp. 127-136). <https://doi.org/10.1145/3560830.3563726>
- [12] Yunus, Y.K.B.M. and Ngah, S.B., **2020**, February. Review of hybrid analysis technique for malware detection. In *IOP Conference Series. Materials Science and Engineering* (Vol. 769, No. 1). IOP Publishing. <https://doi.org/10.1088/1757-899X/769/1/012075>
- [13] Maturure, P., Ali, A. and Gegov, A., **2024**, August. Hybrid machine learning model for phishing detection. In *2024 IEEE 12th International Conference on Intelligent Systems (IS)* (pp. 1-7). IEEE. <https://doi.org/10.1109/IS61756.2024.10705257>
- [14] Ahmad, S., Zaman, M., Al-Shamayleh, A.S., Ahmad, R., Abdulhamid, S.I.M., Ergen, I. and Akhunzada, A., **2024**. Across the spectrum in-depth review AI-based models for phishing detection. *IEEE Open Journal of the Communications Society*, 6, pp.2065-2089. <https://doi.org/10.1109/OJCOMS.2024.3462503>
- [15] Safi, A. and Singh, S., **2023**. A systematic literature review on phishing website detection techniques. *Journal of King Saud University-Computer and Information Sciences*, 35(2), pp.590-611. <https://doi.org/10.1016/j.jksuci.2023.01.004>
- [16] Van Geest, R.J., Cascavilla, G., Hulstijn, J. and Zannone, N., **2024**. The applicability of a hybrid framework for automated phishing detection. *Computers & Security*, 139, p.103736. <https://doi.org/10.1016/j.cose.2024.103736>
- [17] Firdos, R. and Dangi, A., **2026**. SecureScan: An AI-Driven Multi-Layer Framework for Malware and Phishing Detection Using Logistic Regression and Threat Intelligence Integration. *arXiv preprint arXiv:2602.10750*. <https://doi.org/10.48550/arXiv.2602.10750>
- [18] Xia, H., Huang, N., Feng, X., Zhang, R. and Liu, C., **2024**. Starlet: Network defense resource allocation with multi-armed bandits for cloud-edge crowd sensing in IoT. *Digital Communications and Networks*, 10(3), pp.586-596. <https://doi.org/10.1016/j.dcan.2023.03.009>
- [19] Iniyasudhan, S., Kavitha, A., Jeevanth, S., Saravanakumar, S., Ramkumar, A. and Rajesh, T., **2025**, April. I-Driven Solutions to Detect and Prevent Malware and Phishing Attacks. In *International Conference on Web Intelligence and Human-Machine Interaction* (pp. 805-823). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-96-8563-9_52
- [20] Zhao, H., **2022**, April. Target EEG recognition based on CNN and bidirectional GRU hybrid model.

- In *International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2022)* (Vol. 12247, pp. 430-435). SPIE. <https://doi.org/10.1117/12.2636802>
- [21] Khraisat, A., Gondal, I., Vamplew, P. and Kamruzzaman, J., **2019**. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1), pp.1-22. <https://doi.org/10.1186/s42400-019-0038-7>
- [22] Zhang, Q., Zhang, P. and Zhou, F., **2022**. Intraday and interday features in the high-frequency data: Pre-and post-Crisis evidence in China's stock market. *Expert Systems with Applications*, 209, p.118321. <https://doi.org/10.1016/j.eswa.2022.118321>
- [23] Sivasankar, C., Tamilarasan, A., Christy, S. and Parthiban, S., **2025**, May. Towards Practical Phishing Detection: Addressing Challenges with Hybrid Machine Learning Architectures. In *2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)* (pp. 1-4). IEEE. <https://doi.org/10.1109/ICCRTEE64519.2025.11053024>
- [24] Ghafoor, M.I., Marjan, S., Roomi, M.S., Wahid, A., Hussain, S., Baryalai, M. and Alikhail, S., **2022**. Cyber-Malware Defense for Smart Grids Using Machine Learning. *Journal of Applied and Emerging Sciences*, 12(2), pp.190-200. <http://dx.doi.org/10.36785/jaes.122558>
- [25] Jayakeerthi, M., Priya, B.B.V., Santhoshkumar, S.P., Vikram, D. and Chandra, S., **2026**, April. Hybrid CNN–LSTM Framework for Real-Time Detection of Machine Learning-Generated Phishing Attacks. In *2026 Fourth International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)* (pp. 515-521). IEEE. <https://doi.org/10.1109/ICAISS68683.2026.11526143>
- [26] Muhammad, A.U.H., Lee, H. and Choi, J., **2023**. Exploiting mixing regularization for truly unsupervised font synthesis. *Pattern Recognition Letters*, 169, pp.35-42. <https://doi.org/10.1016/j.patrec.2023.03.019>
- [27] Das, B.C., Sartaz, M.S., Reza, S.A., Hossain, A., Nasiruddin, M.D., Bishnu, K.K., Sultana, K.S., Shaty, S.S., Khan, M.D. and Abed, J., **2025**. AI-driven cybersecurity threat detection: Building resilient defense systems using predictive analytics. *arXiv preprint arXiv:2508.01422*. <https://doi.org/10.48550/arXiv.2508.01422>