iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84089

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Machine Learning Framework for Predicting Emerging Cyber Threats

Vasavi Pandi[1], Mr. D. Bhagyaraj Yadav[2]
Department of Computer Science and Engineering Holy Mary Institute of Technology & Science Hyderabad, India

Abstract: ***Modern cyber attacks increasingly involve adaptive adversaries, zero-day exploitation, polymorphic behavior, multi-stage campaigns, and weak signals distributed across heterogeneous security telemetry. Traditional signature-based intrusion detection systems are effective for known threats but provide limited capability for predicting emerging attacks before they fully materialize. This paper presents THREATPREDICT-AI, a machine learning framework for predicting emerging cyber threats through weak-signal anomaly detection, time-series threat forecasting, graph-based threat correlation, ensemble risk scor-ing, explainable prediction artifacts, audit logging, and Human-in-the-Loop analyst validation. The proposed framework ingests safe cyber telemetry such as SIEM logs, IDS alerts, endpoint events, firewall records, authentication logs, and cyber threat intelligence indicators. It converts raw events into event-level, window-level, temporal, and graph-level features. Anomaly risk, forecast risk, and graph risk are fused into a calibrated threat score and mapped to low, medium, or high risk levels. A Flask-based prototype was implemented with a SOC dashboard, role-based access control, synthetic event generation, prediction management, HITL review, configurable model settings, and audit logs. Experimental demonstration using 220 synthetic cyber events generated 59 prediction windows and detected a high-risk emerging threat window with a final score of 0.83. The results show that combining anomaly, temporal, and graph sig-nals improves proactive threat identification while explainability, auditability, and analyst governance improve operational trust.***

Index Terms: ***Cybersecurity, machine learning, emerging cyber threats, anomaly detection, time-series forecasting, graph-based threat intelligence, explainable AI, Human-in-the-Loop, SOC automation, auditability.***

I. INTRODUCTION

The rapid growth of digital infrastructure, cloud services, enterprise networks, Internet of Things devices, and remote access platforms has significantly expanded the cyber attack surface. Modern adversaries do not depend only on simple malware or direct exploitation; they use stealthy, adaptive, and multi-stage strategies that evolve over time. These attacks may involve reconnaissance, credential abuse, privilege escalation, lateral movement, command-and-control communication, and data exfiltration. The early indicators of such attacks are often weak, distributed, and difficult to detect using isolated alerts. Traditional intrusion detection systems (IDS) are largely reactive. Signature-based IDS tools compare observed activity against known attack patterns and indicators of compromise. Although useful for known attacks, these systems cannot reliably detect zero-day threats, polymorphic malware, adver-sarially modified payloads, or new attack behaviors. Anomaly-based IDS methods improve unknown-threat detection, but they frequently suffer from high false positives, concept drift, lack of interpretability, and limited operational trust [1]–[3].

Machine learning offers opportunities to detect complex patterns in security telemetry. However, a single model is rarely sufficient for proactive cyber defense. Emerging threats must be analyzed across behavior, time, and relationships. Behavioral anomaly detection can identify weak deviations. Time-series forecasting can reveal escalation trends. Graph-based modeling can detect coordinated activity among users, IP addresses, assets, event types, and sources. Explainable AI and Human-in-the-Loop (HITL) review are also necessary because SOC analysts must understand and validate high-risk automated predictions before operational response.

This paper proposes THREATPREDICT-AI, a predictive and governance-aware framework for emerging cyber threat detection. The framework combines anomaly risk, forecast risk, and graph risk using an ensemble scoring strategy. It also generates explanation artifacts, stores audit records, and routes high-risk predictions to analysts for validation. The major contributions of this paper are as follows:

1) A unified machine learning framework for predicting emerging cyber threats using heterogeneous cyber teleme-try.
2) An ensemble risk scoring method that combines weak-signal anomaly detection, temporal forecasting, and graph-based threat correlation.

3) An explainability and auditability mechanism that records contributing risk components and analyst decisions.
4) A working Flask-based SOC prototype with role-based access control, dashboards, prediction manage-ment, HITL review, and audit logs.
5) A demonstration using synthetic cyber events showing proactive detection of high-risk windows.

II. RELATED WORK

Machine learning for network intrusion detection has been studied extensively. Sommer and Paxson [1] discussed the challenges of applying machine learning in operational in-trusion detection, including the difficulty of realistic evaluation and deployment outside closed-world settings. Ahmed et al. [2] reviewed network anomaly detection methods and highlighted the importance of handling diverse and evolving traffic behavior. Chandola et al. [3] provided a broad survey of anomaly detection techniques and their application domains.

Deep learning has also been applied to intrusion detection and malware analysis. LSTM networks are useful for modeling sequential dependencies and long-range temporal behavior [4]. Time-series models, including classical forecasting techniques, can support early warning when threat intensity changes gradually [5]. However, purely temporal models may miss structural relationships among cyber entities.

Graph-based cyber analytics represent users, hosts, IP ad-dresses, processes, vulnerabilities, and indicators as nodes, and communications or co-occurrences as edges. Such models can capture lateral movement, coordinated campaigns, and entity relationships that are not visible in isolated event records [6]. Cyber threat knowledge bases such as MITRE ATT&CK provide structured representations of adversarial tactics and techniques, supporting contextual analysis and threat reasoning[7].

Explainable AI is important in security operations because analysts require evidence behind automated predictions. LIME and SHAP are widely used model-agnostic explanation tech-niques for interpreting machine learning outputs [8], [9]. HITL systems further improve trust by allowing analysts to validate, reject, or annotate high-risk predictions [10]. Gov-ernance frameworks such as the NIST AI Risk Management Framework emphasize transparency, accountability, and risk control for AI-enabled systems [11].

Existing studies often address intrusion detection, anomaly detection, forecasting, graph analytics, explainability, or hu-man oversight separately. The proposed framework integrates these components into a single practical architecture for proac- tive cyber threat prediction.

III. PROPOSED FRAMEWORK

The proposed THREATPREDICT-AI framework is designed as a defensive cybersecurity analytics system. It does not execute attacks, exploits, malware, or unauthorized scanning. It uses safe telemetry and synthetic events to demonstrate predictive cyber defense. The framework consists of five major layers: data ingestion, preprocessing and feature engineering, analytical modeling, decision fusion and explainability, and governance with SOC presentation. Fig. 1 shows the overall architecture. The ingestion layer receives SIEM logs, IDS/IPS alerts, endpoint telemetry, fire-wall events, authentication logs, and CTI indicators. The preprocessing layer normalizes heterogeneous event fields and handles missing values. The feature engineering layer creates event-level and window-level features. The analytical layer computes anomaly risk, forecast risk, and graph risk. The decision layer fuses these signals and generates explanations. The governance layer supports HITL review and audit logging.

IV. METHODOLOGY

Let E denote the complete cyber event stream. For a time window t, the batch of events is represented as

$$B_t = \{e_1, e_2, \ldots, e_{nt}\}, \quad (1)$$

where n_t is the number of events observed in the window. Each event is converted into a feature vector:

$$\phi(e_i) = [x_{i1}, x_{i2}, \ldots, x_{id}], \quad (2)$$

where d is the number of extracted features. The window feature matrix is

$$X_t \in R^{nt \times d}. \quad (3)$$

A. *Weak-Signal Anomaly Detection*

The anomaly module estimates whether a window deviates from expected behavior. Let $\alpha(\phi(e_i))$ be the anomaly score for event e_i. The window-level anomaly risk is computed as

$$A_t = \frac{1}{|B_t|} \sum_{e_i \in B_t} \alpha(\phi(e_i)). \quad (4)$$

In the prototype, Isolation Forest is used when sufficient data are available. A severity and suspicious-event fallback score is used when the number of windows is small.

B. Time-Series Threat Forecasting

Threat intensity is computed as the sum of event severities in the window:

$$I_t = \sum_{e_i \in B_t} sev(e_i). \quad (5)$$

Given historical intensities $\{I_{t-k}, \ldots, I_t\}$, the forecasting module estimates future intensity $\hat{I}_{t+1:t+h}$. The forecast-based

$$F_t = g\ \mu(I_{t+1:t+h}),\ \sigma(I_{t+1:t+h}) \quad , \quad (6)$$

where μ and σ denote the mean and standard deviation of predicted intensity and $g(\cdot)$ maps the result to [0, 1].

C. Graph-Based Threat Correlation

For each window, a threat graph is constructed as

$$G_t = (V_t, E_t), \quad (7)$$

where nodes represent users, assets, IP addresses, sources, event types, and indicators, while edges represent co-occurrence or communication relationships. Graph density is

$$\text{Density}(G_t) = \frac{|E_t|}{|V_t|(|V_t| - 1)/2}. \quad (8)$$

and hubness is

$$\text{Hubness}(G_t) = \frac{\max_{v \in V_t} \deg(v)}{|V_t| - 1}. \quad (9)$$

$G^{(r)} = \beta_1 \text{Density}(G_t) + \beta_2 \text{Hubness}(G_t) + \beta_3 C(G_t)$, (10) where $C(G_t)$ represents graph cohesion.

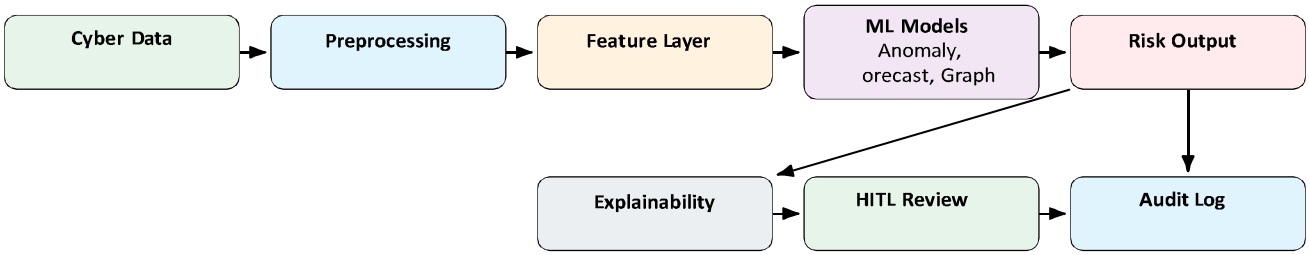


Fig. 1. Overall architecture of the proposed THREATPREDICT-AI framework.

TABLE I
MAJOR IMPLEMENTATION MODULES

Module	Function
Ingestion	Stores logs, alerts, authentication records, and CTI events.
Preprocessing	Cleans data, normalizes severity, parses time, and encodes categories.
Feature Service	Builds event-level, window-level, and graph-level features.
Anomaly Service	Computes weak-signal anomaly risk.
Forecasting Service	Estimates near-term threat intensity.
Graph Service	Constructs entity graphs and computes structural metrics.
Ensemble Service	Produces calibrated final risk and risk level.
Explainability	Generates evidence, rationale, and recommendations.
HITL Review	Allows analyst approval, rejection, or downgrade.
Audit Service	Records user actions, predictions, and review decisions.

D. Ensemble Risk Prediction

The final risk score combines anomaly, forecast, and graph risk:

$$r_t = \sigma\left(\gamma\left(w_A A_t + w_F F_t + w_G G_t^{(r)} - 0.5\right)\right), \quad (11)$$

where $\sigma(x) = 1/(1 + e^{-x})$, γ controls calibration sharpness, and $w_A + w_F + w_G = 1$. The risk level is assigned as

$$\ell_t = \begin{cases} \text{HIGH}, & r_t \geq \vartheta_H \\ \text{MEDIUM}, & \vartheta_M \leq r_t < \vartheta_H \\ \text{LOW}, & r_t < \vartheta_M \end{cases} \quad (12)$$

TABLE II
MODEL CONFIGURATION USED IN DEMONSTRATION

Parameter	Value
Anomaly weight w_A	0.40
Forecast weight w_F	0.30
Graph weight w_G	0.30
Calibration factor γ	5.0
High threshold ϑ_H	0.75
Medium threshold ϑ_M	0.45

TABLE III
OBSERVED SYSTEM RESULTS

Metric	Observed Value
Synthetic cyber events	220
Prediction windows	59
High-risk predictions	1
Pending HITL reviews	1
High-risk final score	0.83
High-risk anomaly risk	1.00
High-risk forecast risk	0.89
High-risk graph risk	0.52
Mean severity in high-risk window	8.0

V. RESULTS AND EVALUATION

The system was evaluated using safe synthetic cyber teleme-try generated by the prototype. The data simulator produced normal events, suspicious events, failed logins, malware alerts, port scans, data transfers, and high-risk clustered patterns. The experiment generated 220 synthetic cyber events and produced 59 prediction windows. The high-risk prediction was generated when a window showed three high-severity events, strong anomaly risk, high forecast risk, and meaningful graph correlation risk. The dominant contributing component was anomaly risk, indicating that the observed window strongly deviated from expected behavior. The dashboard also showed the risk distribution and threat intensity trend, while the HITL page routed high-risk predictions to the analyst for validation.

The results demonstrate that the framework can distinguish between low, medium, and high threat windows based on combined risk evidence. Windows with moderate anomaly and forecast values are classified as medium risk, while lower ensemble scores are triaged as low risk. The high-risk case is routed to HITL review, supporting governance and operational control.

TABLE IV
REPRESENTATIVE PREDICTION WINDOWS

ID	A_t	F_t	G_t	r_t	Risk
59	0.81	0.64	0.28	0.62	Medium
58	0.21	0.65	0.37	0.37	Low
57	0.44	0.84	0.41	0.56	Medium
56	0.20	0.71	0.37	0.38	Low
55	0.25	0.69	0.34	0.39	Low
54	0.51	0.72	0.43	0.56	Medium
53	0.52	0.86	0.41	0.61	Medium
52	0.21	0.77	0.36	0.41	Low

VI. DISCUSSION

The proposed framework improves over traditional IDS in four ways. First, it moves from reactive signature matching to proactive threat prediction. Second, it combines multiple analytical perspectives rather than relying on a single model. Third, it provides explanation artifacts that connect risk predictions to observable evidence. Fourth, it includes HITL review and audit logging, which are necessary for SOC trust and governance.

Although the demonstration uses synthetic data, the proto-type validates the feasibility of integrating anomaly detection, temporal forecasting, graph correlation, ensemble scoring, dashboard visualization, and analyst review in a single system. The framework is especially useful for academic demon-stration and can be extended toward real SIEM integration, streaming ingestion, advanced deep learning, graph neural networks, and model drift monitoring.

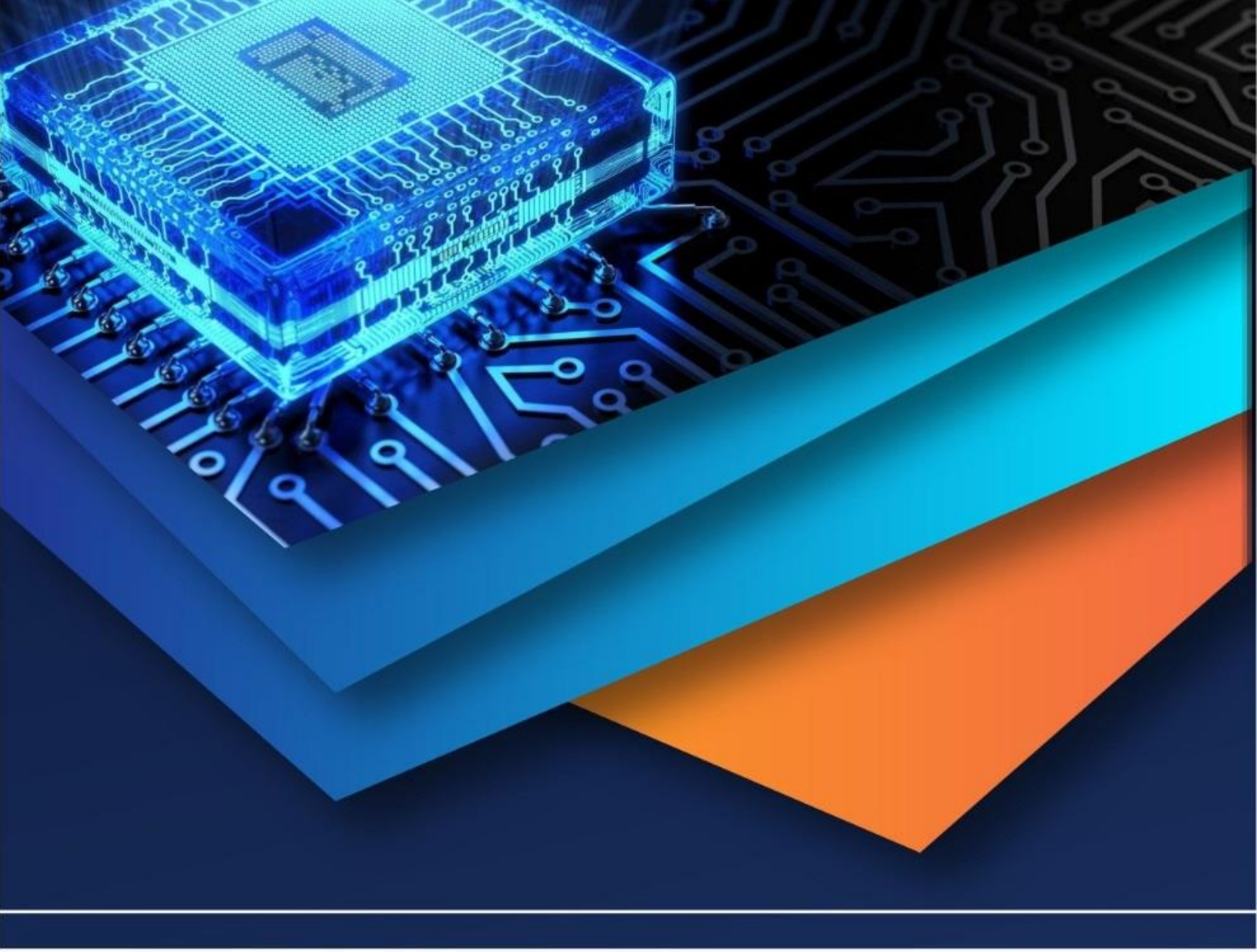
VII. CONCLUSION

This paper presented THREATPREDICT-AI, a machine learning framework for predicting emerging cyber threats. The framework integrates heterogeneous cyber event ingestion, preprocessing, weak-signal anomaly detection, time-series forecasting, graph-based threat correlation, ensemble prediction, explainability, auditability, and Human-in-the-Loop validation. A Flask-based prototype demonstrated the end-to-end workflow using synthetic cyber events. The system generated 59 prediction windows from 220 events and detected a high-risk window with a final risk score of 0.83. The results show that integrated anomaly, temporal, and graph-based risk analysis can support proactive cyber defense while explainability, audit logging, and analyst validation improve operational trust. Future work includes evaluating the framework on larger real-world datasets, integrating streaming SIEM data, using LSTM or Transformer-based forecasting, applying graph neural networks, adding SHAP-based explanations, implementing adversarial robustness tests, and incorporating feedback-driven model adaptation from analyst decisions.

REFERENCES

[1] R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in Proc. IEEE Symposium on Security and Privacy, 2010, pp. 305–316.
[2] M. Ahmed, A. N. Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” Journal of Network and Computer Applications, vol. 60, pp. 19–31, 2016.

[3] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.
[4] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[5] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Wiley, 2015.
[6] W. Wang and Z. Lu, “A survey on graph-based cyber security analytics,” IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 1232–1263, 2021.
[7] MITRE Corporation, “MITRE ATT&CK framework,” 2024. [Online]. Available: https://attack.mitre.org
[8] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” in Proc. ACM SIGKDD, 2016, pp. 1135–1144.
[9] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Advances in Neural Information Processing Systems, 2017.
[10] S. Amershi, M. Cakmak, W. B. Knox, and T. Kulesza, “Power to the people: The role of humans in interactive machine learning,” AI Magazine, vol. 35, no. 4, pp. 105–120, 2014.
[11] National Institute of Standards and Technology, “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY