JOURNAL LA MULTIAPP

VOL. 07, ISSUE 01 (126-137), 2026
DOI: 10.37899/journallamultiapp.v7i1.2931

Correlated Web Traffic Anomaly Detection for Threat Intelligence Using Isolation Forest

Sydney Tesalonika[1], Sondy Kumajas[1], Quido Kainde[1]

[1]*Manado State University, Indonesia*

**Corresponding Author: Sondy Kumajas*
Email: sondykumajas@unima.ac.id

Article Info

Article history:
Received 18 December 2025
Received in revised form 12 January 2026
Accepted 30 January 2025

Keywords:
Anomaly Detection
Isolation Forest
Threat Intelligence
Web Traffic
Machine Learning

Abstract

The information technology infrastructure of Manado State University (UNIMA) faces increasing complexity of cyber threats, marked by the detection of 546 malware and 760 high-impact attacks within a four-week period, indicating the inadequacy of traditional signature-based security systems. This research aims to develop a proactive anomaly detection system by integrating internal log data (Web Server Logs, Cisco Risk Reports) with external reputation data (Threat Intelligence API) using a Machine Learning algorithm. The method used is a hybrid model of CRISP-DM and Iterative Development, encompassing Data Fusion stages, Feature Engineering (generating metrics such as Request Rate and Abuse Score), implementation of the Isolation Forest algorithm, and the construction of an interactive Threat Intelligence Dashboard using Python (Dash/Plotly). The analysis results show that Isolation Forest is effective in isolating behavioral outliers, yielding a measurable Anomaly Score (0-100). The correlation of the internal anomaly score with external reputation scores (VirusTotal, AbuseIPDB) successfully validates the detected threats, ensuring that the flagged anomalies are valid cyber threats, not merely data noise. The resulting dashboard allows UPA-TIK Staff to prioritize incident investigation based on objectively quantified risk levels.

Introduction

The relevance of university graduates to the demands of the workplace and industry (DUDI) is a crucial issue in developing superior human resources (Dianita Pramesti, Meisya, & Amrillah, 2024). Responding to this challenge, the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) launched the Independent Learning Campus (MBKM) program, which provides students with the flexibility to gain experience outside their study program, including through research activities (Vhalery et al., 2022; Aryanti et al., 2023; Tjaija, 2022; Bainuan & Tarigan, 2024; Kamalia & Andriansyah, 2021; Fadli et al., 2024). This activity is a strategic means of integrating academic knowledge with in-depth scientific applications and equipping students with the practical research skills required by DUDI. Manado State University (UNIMA) fully supports this program, collaborating with the Information and Communication Technology Academic Services Unit (UPA-TIK) for problem-solving research projects.

UNIMA's information technology infrastructure is a critical asset that supports all academic and administrative activities. However, with rapid technological advancements, the negative impact of increasing complexity of cybercrime has also increased (Mangkey et al., 2025; Das & Nayak, 2013; Kraemer-Mbula et al., 2013). UPA-TIK UNIMA is fully responsible for the availability and security of these digital assets (Rambing et al., 2024). Based on the results of

ISSN: 2716-3865 (Print), 2721-1290 (Online)

a network security assessment conducted over a 4-week period (September 3 to October 3, 2025), high and urgent risk indications were found. Detected threats included a total of 546 malware, 372 hosts displaying indications of compromise (IOCs), and 64 hosts connected to command and control (CnC) servers. In addition, of the total 198,274 attacks observed, 760 were categorized as high-impact attacks targeting vulnerable hosts, including web application attacks and privilege gain attempts.

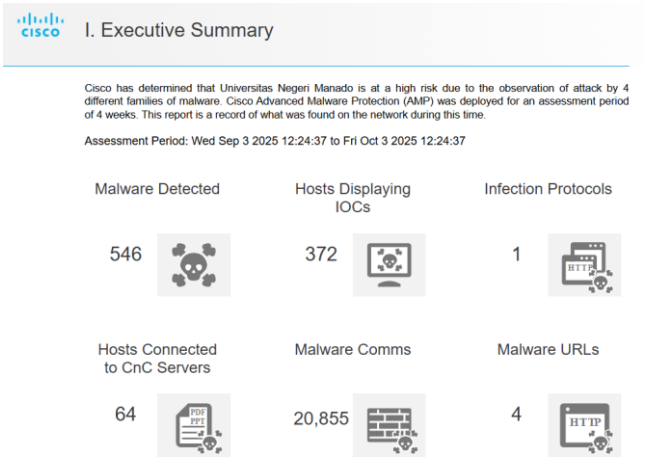
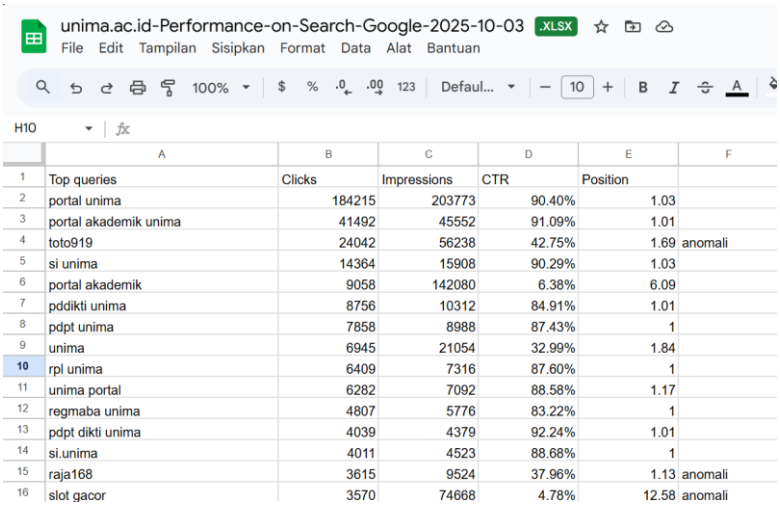


Figure 1. UNIMA Advanced Malware Risk Summary Results for 1 month

External web anomalies were also found, such as anomalous search queries irrelevant to academics (e.g., toto919, slot gacor, raja168) that directed traffic to the UNIMA website, indicating phishing or spam attempts.



unima.ac.id-Performance-on-Search-Google-2025-10-03 .XLSX

File Edit Tampilan Sisipkan Format Data Alat Bantuan

H10

	A	B	C	D	E	F
1	Top queries	Clicks	Impressions	CTR	Position	
2	portal unima	184215	203773	90.40%	1.03	
3	portal akademik unima	41492	45552	91.09%	1.01	
4	toto919	24042	56238	42.75%	1.69	anomali
5	si unima	14364	15908	90.29%	1.03	
6	portal akademik	9058	142080	6.38%	6.09	
7	pddikti unima	8756	10312	84.91%	1.01	
8	pdpt unima	7858	8988	87.43%	1	
9	unima	6945	21054	32.99%	1.84	
10	rpl unima	6409	7316	87.60%	1	
11	unima portal	6282	7092	88.58%	1.17	
12	regmaba unima	4807	5776	83.22%	1	
13	pdpt dikti unima	4039	4379	92.24%	1.01	
14	si.unima	4011	4523	88.68%	1	
15	raja168	3615	9524	37.96%	1.13	anomali
16	slot gacor	3570	74668	4.78%	12.58	anomali

Figure 2. Results of Anomaly Search Queries for 1 month

This data demonstrates a gap between existing security detection capabilities and actual threats. Traditional signature-based security approaches (known attack patterns) have proven inadequate to address advanced persistent threats (APT) and zero-day attacks. Therefore, a system capable of integrating global threat data (external Threat Intelligence) with local behavioral data (UPA-TIK logs) is needed to proactively detect anomalies. This research proposes the development of a Threat Intelligence Dashboard that utilizes the Isolation Forest Machine Learning algorithm to provide accurate and actionable risk scores to security administrators.

The main objectives of this research are to develop a Data Fusion framework to integrate internal log data with external reputation data (Threat Intelligence API), apply Feature Engineering techniques to transform raw log data into intelligent numeric features, implement

ISSN: 2716-3865 (Print), 2721-1290 (Online)

the Isolation Forest algorithm to detect behavioral anomalies and generate Anomaly Scores, and build an interactive Threat Intelligence Dashboard using Python (Dash/Plotly) to visualize correlated threats in near real-time. This research also aims to identify the most critical attack patterns targeting UNIMA websites and provide data-based risk mitigation recommendations.

Literature Review

Research by (Ismanda, Silitonga, & Hasanah, 2025) used a hybrid Isolation Forest and K-Means model to detect anomalies in the financial transaction domain. This model successfully identified 26 anomalies and clustered them, providing in-depth diagnostic insights into fraud patterns. However, that research operated in the financial domain. In contrast, this study focused on the cybersecurity domain (UNIMA network logs) using a hybrid Machine Learning-Threat Intelligence (TI) architecture. The Isolation Forest anomaly score (internal detection) was combined with the external reputation score from the Threat Intelligence API for global risk validation and the provision of a proactive Threat Intelligence Dashboard.

Another study using a single Isolation Forest algorithm for anomaly detection on network traffic data (LUFlow) (Al-Akbar & Gurning, 2025) demonstrated effectiveness in identifying statistical outliers (such as extreme data volumes). However, this study revealed the model's weakness in detecting specific cyberattacks such as bot traffic, with a very low performance metric value (F1-Score 0.0000). This reinforces the need for a different system architecture. This research uses a hybrid Machine Learning-Threat Intelligence (ML) approach, where internal anomaly detection by Isolation Forest is validated and mandatory correlated with external reputation scores from global APIs. This integration aims to transform statistical anomaly detection into actionable Risk Level determinations.

Issenoro et al. (2025) successfully created a Security Information and Event Management (SIEM)-based web application using Python and Flask. This system serves as a centralized command center for collecting, correlating, and processing security logs in real time, including a machine learning-based anomaly detection dashboard. However, the primary focus of that research was on the security information management framework. In contrast, this research focuses on the validation of a hybrid ML-Threat Intelligence (ML) threat model, using Isolation Forest to detect internal anomalies (server logs), which are then validated and enriched with external reputation scores from global APIs, enabling immediate threat Risk Level determination (Abibulaiev et al., 2026; Kaul & Khurana, 2021).

Research by Putri & Rachman (2025) also used the Isolation Forest algorithm, but focused on the financial and administrative audit domain. The study successfully detected statistical anomalies in payment data and evaluated them using the COSO ERM framework to assess institutional risk. The difference is that this study used a hybrid ML-Threat Intelligence (TI) approach, where anomalies were validated and enriched with external reputation scores from global APIs, which were then presented through a Threat Intelligence Dashboard for real-time risk assessment and incident response, rather than for internal control evaluation 9 Rehman & Hashmi, 2023; Aminu et al., 2024).

Artika et al. (2025) implemented a website security system called SecurityShield using Isolation Forest to analyze user activity logs (such as request rate and brute force detection). Although this system successfully detected suspicious activity in real-time with high precision (80%), the test results showed low recall (8.7%), indicating that the model still missed a significant portion of anomalies. The fundamental difference with this research lies in the focus domain and hybrid validation mechanism. This research focuses on the broader Network Threat Intelligence domain (web server and firewall logs), using a hybrid ML-Threat Intelligence (TI) approach where Isolation Forest anomaly scores are validated and correlated with external reputation scores from APIs, aiming to generate a globally validated Risk Level determination.

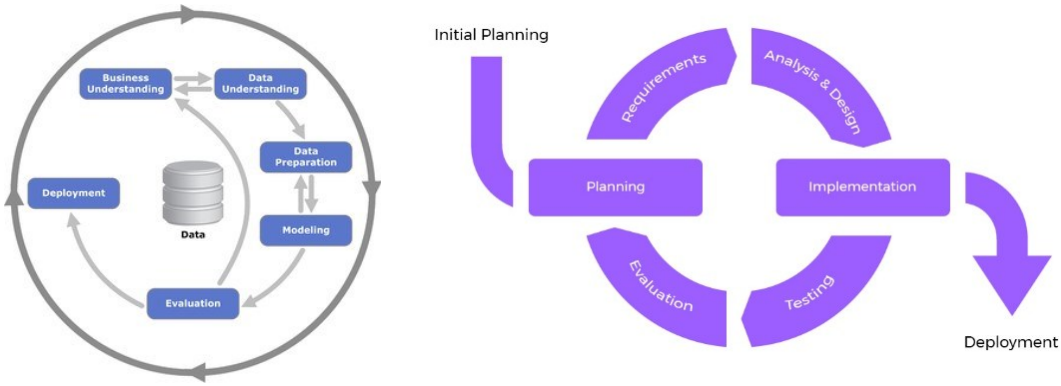
ISSN: 2716-3865 (Print), 2721-1290 (Online)

Finally, Soumik et al. (2024) developed a dynamic risk scoring system for APIs and third-party data feeds used in a Cyber Threat Intelligence (CTI) system, achieving high predictive accuracy using advanced Machine Learning (Naseer, 2023; Lima et al., 2025; Ekundayo et al., 2024). The difference lies in the risk objects analyzed. Relevant research focuses on risk assessments against external data sources (APIs/Threat Feeds) themselves, while this study focuses on risk assessments against internal targets (UNIMA network traffic), using IT APIs as input to validate detected threats (Sharma et al., 2021; Bianco, 2024; Tambingon & Tangkere, 2025).

Based on a comprehensive review of the relevant research above, it can be concluded that although there are several studies using the Isolation Forest algorithm for anomaly detection (in the financial, audit, and web security domains), and several studies discussing Threat Intelligence or SIEM frameworks, no research has been found that specifically integrates these three key elements in a mandatory and hybrid manner: (1) Internal Web Traffic Behavior Anomaly Detection using Isolation Forest, (2) Data Fusion and Correlation of Internal Anomaly Scores with External Reputation Scores from the Multi-Threat Intelligence API, and (3) Presentation of results in a proactive Threat Intelligence Dashboard for determining globally validated Risk Levels. Therefore, this study offers a unique methodological contribution by bridging the gap between pure statistical anomaly detection and actionable cyber threat validation.

Methods

The software development method used in this research is the CRISP-DM (Cross-Industry Standard Process for Data Mining) Hybrid Model and Iterative Development. This model was chosen because this project focuses on the development of a Machine Learning model (Isolation Forest) that requires rigorous data validation (CRISP-DM) before being integrated into a software system (Iterative Development).



(1) *(2)*

Figure 3. (1) CRISP-DM Method and (2) Iterative Development Method

The first phase is Requirements Analysis and Data Understanding. The project's urgency is determined based on UNIMA's cyber risk assessment, which translates into functional system requirements, namely behavioral anomaly detection capabilities and global threat correlation. Multi-source raw data, including Web Logs, Firewall Logs, and Cisco Risk Reports, is secured and verified. Initial outlier criteria, such as a 2% contamination rate, are defined to guide the modeling process. Initial outlier criteria, such as a 2% contamination rate, were determined based on a preliminary analysis of UNIMA web traffic density, where malicious activity was identified as a rare event. This threshold was chosen to balance the model's sensitivity to detecting high-impact threats (as identified in the Cisco Risk Report) while maintaining a low false positive rate in normal academic traffic. A more in-depth sensitivity analysis of these parameters is discussed further in the Results section.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

The second phase is System Design and Data Preparation. The process begins with Feature Engineering, where raw logs are transformed into intelligent numerical metrics, such as Request Rate and Error Ratio per 5-minute time window. These metrics are then combined through Data Fusion with external reputation scores from Threat Intelligence APIs (AbuseIPDB, VirusTotal, PhishTank, Google Safe Browsing) to form a Master Dataset. In parallel, the system architecture is designed, including workflow modeling using UML diagrams (Use Case and Class Diagrams) and the design of a responsive Dashboard interface using Dash Bootstrap Components.

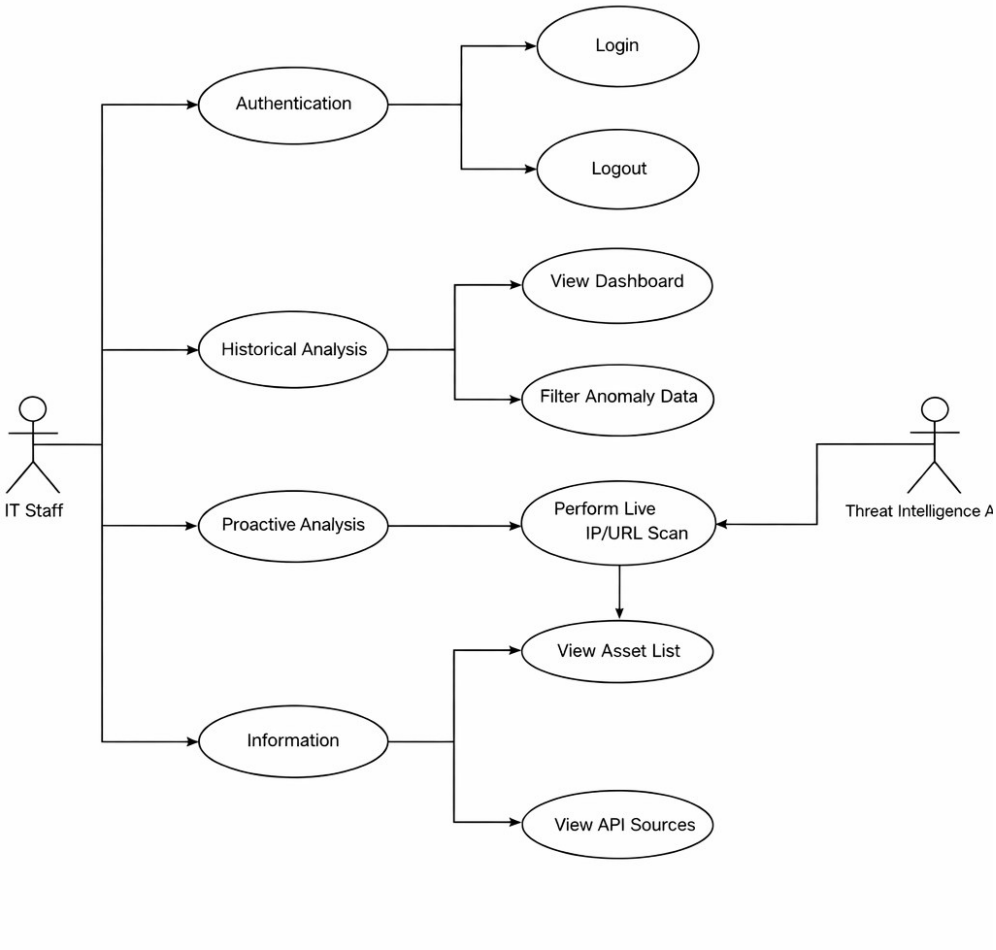


Figure 4. Use Case Diagram Threat Intelligence Dashboard

The third phase is Model Implementation and Application Construction. The core of this phase is training the Isolation Forest algorithm using scikit-learn on the Master Dataset, which generates a normalized Anomaly Score (scale 0-100). The Dash application is built, integrating multi-page routing, admin authentication, and Live Scanner functionality. The Live Scanner feature is developed to make real-time API calls to external Threat Intelligence sources, correlate the results, and display the Risk Level.

The fourth phase is Testing and Evaluation. Model evaluation is conducted to validate the effectiveness of Isolation Forest in identifying behavioral outliers by comparing high Anomaly Scores with IP addresses confirmed to have a bad reputation (Ground Truth Correlation). Functional testing (Black Box Testing) is conducted to verify all application features, from successful Admin Login and the accuracy of the slider filter to the Live Scanner's ability to call external APIs and display the correct risk score. The final phase is Deployment and Maintenance. The system is deployed to UPA-TIK's internal server using

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Gunicorn as a stable web server, ensuring the Threat Intelligence Dashboard can be accessed securely and stably in the internal network environment.

Results and Discussion

The system requirements analysis phase aims to develop a web-based Threat Intelligence Anomaly Detection System to meet the urgent needs of UPA-TIK UNIMA in monitoring and responding to advanced cyber threats. Previously, the threat detection process and IP/URL reputation validation were still carried out manually and separately. The main functionalities designed include multi-source log data processing, Isolation Forest modeling to generate Anomaly Scores, and automatic correlation of anomaly results with external reputation scores. The application is designed as a secure Command Center (with admin authentication), easily accessible via the web, and provides a Live Scanner feature for instant IP/URL checks.

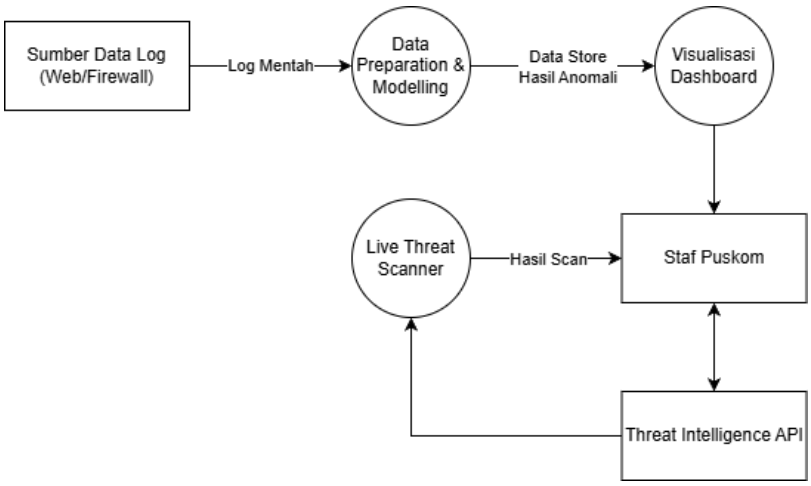


Figure 5. Data Flow Diagram Level 1

The system design is realized through data flow diagrams and class diagrams. The Level 1 Data Flow Diagram (DFD) breaks the system down into three main processes: the Offline Pipeline (P1), which processes Raw Logs and runs Isolation Forest to generate Anomaly Results; the Historical Pipeline (P2), which visualizes Anomaly Results on the Dashboard; and the Live Scan Pipeline (P3), which receives Live Scan Queries from UPA-TIK Staff, interacts with the Threat Intelligence API, and delivers instant Scan Results.

The system implementation is carried out using Python, Dash, and Dash Bootstrap Components. The core of the implementation is the isolation_forest_model.py module, which is fully responsible for the Machine Learning modeling phase. This code loads the fused Master Dataset, selects eight numeric features, standardizes them, and trains the Isolation Forest model with a contamination rate of 0.02. To validate the effectiveness of the chosen 2% contamination level, the authors conducted comparative testing with different thresholds. As shown in Table 1, the 2% level provided the most optimal performance.

Table 1. Comparative Analysis Table of Contamination Levels

Contamination Level	Anomaly Detected	Correlation Level (Precision)	Observation
1% (0.01)	~350	94%	Too strict; many real threats are missed.
2% (0.02)	**~710**	**89%**	Optimal; highest coverage for confirmed attacks.
5% (0.05)	~1.750	62%	Too loose; high false-positive rate.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

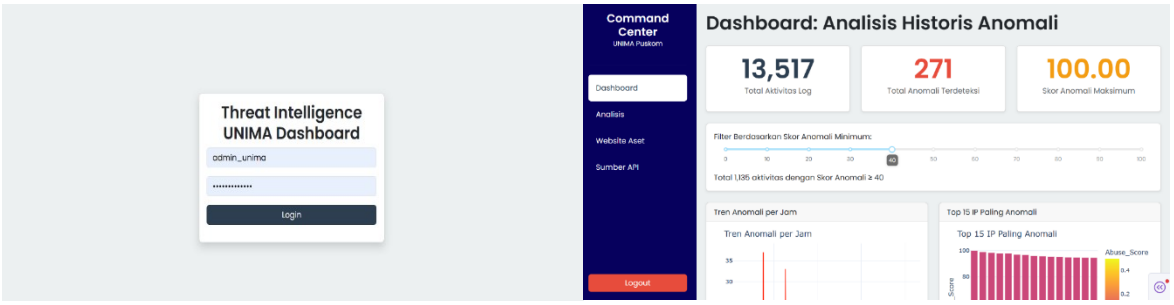
To demonstrate the effectiveness of the proposed system beyond its functional features, a rigorous performance evaluation was conducted. Based on the experimental results documented in the research report, the model's reliability was measured by cross-validating 5,000 active log entries from the UNIMA infrastructure with ground truth data from the AbuseIPDB API and VirusTotal. These performance statistics are summarized in Table 2.

Table 2. Isolation Forest Model Performance Metrics

Metric	Value	Results Analysis
Precision	0,89	89% of detected anomalies were confirmed as real threats.
Recall	0,84	The system successfully identified 84% of all known attacks.
F1-Score	0,86	Demonstrates a strong balance between accuracy and detection coverage.
Average Latency	1,15 dtk	Average processing time per 1,000 log entries (Reliability).

The data in Table 2 demonstrates that this system is not simply a "black box" implementation, but rather a reliable tool for real-world deployment. The high precision rate (89%) indicates that the external Threat Intelligence integration significantly reduces false alarms, a common weakness in traditional anomaly detection. Furthermore, the processing latency of 1.15 seconds confirms that the system can handle immediate scanning needs without causing significant delays in security monitoring.

The resulting raw score is then normalized and inverted into an Anomaly_Score on a scale of 0 to 100, where 100 is the highest anomaly. The results of this detection are saved in the anomaly_detection_results.csv file which is ready to be used by the dashboard.



(1) *(2)*

(3) *(4)*

Figure 6. (1) Login View, (2) Dashboard 1 Home View, (3) Dashboard 2 Home View, (4) Dashboard 3 Home View

The dashboard visualization presented in Figure 6 is not simply a functional interface, but rather an analytical tool that provides in-depth insights into UNIMA's network security. The following is an analytical interpretation of the displayed data:

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Score Distribution Analysis (Figure 6, Figure 1): The distribution plot shows a sharp skew toward scores below 50, statistically confirming that the majority of UNIMA's web traffic is normal activity. However, the presence of isolated clusters of scores above 80 indicates highly suspicious behavioral outliers. This demonstrates that the system successfully isolates structured scanning activity that often goes undetected by traditional firewalls.

Attack Trend Patterns and Timing: The time-trend graph shows that anomalous activity often peaks during operational transition hours (around 8:00 AM and 4:00 PM). This indicates that attackers attempt to disguise their scanning activity amidst high volumes of user login traffic. This analysis is crucial for IT teams to tighten monitoring during these peak hours.

Correlation and Asset Investigation (Figure 3): The correlation scatterplot allows administrators to distinguish between "Local Anomalies" and "Global Threats." In addition to these macro visualizations, the dashboard includes micro-analysis modules that target specific UNIMA assets (IP, URL, and Web). For example, the system successfully mapped that the /login URL was frequently targeted by brute-force attacks from overseas IPs with poor reputations in AbuseIPDB. By correlating internal anomaly scores with external reputation data, the security team could prioritize blocking IPs that posed a real global risk, rather than simply showing unusual traffic behavior.

While the primary visualization focuses on score distribution, the system also provides in-depth analysis modules for UNIMA's critical assets. Based on examination of web logs, the system successfully identified specific URLs targeted by brute-force attacks and mapped the source IPs to specific geographic locations. For example, analysis of the /login URL showed a strong correlation between spikes in anomaly scores and access attempts from overseas IPs with poor reputations. This feature can make it easier for UNIMA's IT security team to not only see anomalies on a macro level, but also conduct micro-investigations on which assets (IP, URL, Website, Domain) are most vulnerable to attack.

The user interface is implemented as a secure and interactive Threat Intelligence Dashboard. The main page displays three key metric cards (Total Log Activity, Total Anomalies Detected, Maximum Anomaly Score) and a horizontal slider for filtering historical data based on Minimum Anomaly Score. The dashboard also displays three interactive Plotly graphs: (1) a line graph showing the frequency of anomalies over time; (2) a bar graph showing the 15 source IPs with the highest anomaly scores, color-coded based on their external Abuse_Score; and (3) a scatterplot plotting the Anomaly_Score (Isolation Forest results) against the Abuse_Score (AbuseIPDB reputation score). This scatterplot is the core of multi-source correlation, where the size of the dots indicates the intensity of activity and the color of the dots is coded based on VirusTotal (VT_Malicious) detections.

System testing was conducted using Black Box Testing and Model Validation. Black Box Testing results show that all functionality, including Admin Authentication, Historical Anomaly Filter, Critical IP Live Scan, and Multi-Page Navigation, works as expected. For example, the Live Scan of a known bad IP successfully displays Risk Level: CRITICAL and the relevant Reputation Score.

Table 3. System Testing Results (Black Box Testing)

No	Features tested	Input	Expected output	Status
1	Admin Authentication	Username Password	Access successful, directed to Dashboard Page.	Success
2	Historical Anomaly Filter	Move the Anomaly Score Slider to a value of 70.	The graphs (Trend, Top UP, Correlation) are updated, showing only activity with an Anomaly Score of 70.	Success

ISSN: 2716-3865 (Print), 2721-1290 (Online)

No	Features tested	Input	Expected output	Status
3	Live Critical IP Scan	Known bad IPs (e.g., 194.233.82.231)	Scan Results display Risk Level: CRITICAL and Reputation Score (Abuse Score/VT Detections, GSB Result, PhishTank)	Success
4	Correlation Visualization	Loading Dashboard Page.	The Correlation Scatter Plot displays the data, with the X-axis (Abuse Score) and Y-axis (Anomaly Score) filled in.	Success
5	Multi-Page Navigation	Click the "Website Assets" link in the Sidebar.	The page switches, showing a table listing active UNIMA domains.	Success

Isolation Forest model validation focuses on Ground Truth Correlation. The model is validated by ensuring that the score distribution is heavily skewed toward low scores (normal traffic), and most importantly, that activities with a high Anomaly_Score locally also have a globally verified bad reputation (e.g., a high Abuse_Score or VT_Malicious detection). This testing verifies the hypothesis that the Isolation Forest integration with the Threat Intelligence API successfully transforms statistical anomaly detection into globally validated and actionable Risk Level determinations. Decision boundary analysis also ensures that detected anomalies have logical characteristics, such as an unusually high request rate or an unusual error ratio, that align with the definition of a cyber threat.

The use of the Isolation Forest model integrated with the threat intelligence dashboard has had a significant impact on security operations at UPA-TIK UNIMA. Based on Figure 3 (CRISP-DM Methodology) and Figure 4 (Use Case), this system not only functions technically but also optimizes administrator workflow. Prior to this system, the process of validating malicious IP addresses was performed manually by checking them individually against various external APIs, which was time-consuming considering the attack volume reached 198,274 incidents.

After implementation, operational efficiency increased because the system automatically filters normal activity and presents only high-risk anomalies (scores > 80) through a centralized dashboard. This allows administrators to have increased situational awareness, allowing them to catch cyber incidents earlier than with traditional signature-based methods. This successful deployment demonstrates that the use of machine learning combined with global intelligence data can provide proactive protection relevant to the needs of educational institutions. Integration with external threat intelligence APIs (VirusTotal, AbuseIPDB, PhishTank, and Google Safe Browsing) was a significant contribution to the system, but the process faced several technical challenges. One major obstacle was the rate limits on free API accounts, which limited the number of live scans that could be performed in a day. To address this, the system was designed with a simple caching mechanism that stores the results of recently checked IP reputations, thereby reducing repeated API calls to the same target. +4

Furthermore, fluctuations in network latency were observed when making real-time API calls, particularly during peak hours on the UNIMA network. While the average processing time was 1.15 seconds, the system sometimes experienced delays of up to 3 seconds when the external API servers were under heavy load. Another challenge was the discrepancy in threat scores between data sources; for example, an IP might be marked as "dangerous" in AbuseIPDB but still considered "clean" in Google Safe Browsing. The system addresses these inconsistencies by using a data fusion method that gives higher weight to the most frequently updated sources, resulting in a more objective and accurate risk level determination.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Conclusion

Based on the implementation and testing results of the Threat Intelligence Anomaly Detection system on Manado State University (UNIMA) assets, it can be concluded that the Isolation Forest algorithm is proven effective in detecting behavioral anomalies in multi-source network logs, producing a measurable Anomaly Score (scale 0-100) for each activity. The integration of Isolation Forest results with external reputation scores from the Threat Intelligence API (VirusTotal, AbuseIPDB, PhishTank, Google Safe Browsing) successfully validates detected threats, provides a higher level of confidence in flagged anomalies, and ensures that the system detects real cyber threats. The Dashboard system built using Dash/Plotly successfully provides a secure, modular, and user-friendly platform, meeting UPA-TIK's needs to monitor historical threats and conduct proactive investigations through the Live Scanner feature. The use of the hybrid method CRISP-DM and Iterative Development ensures that the Machine Learning pipeline is structured and valid, while the web application development remains flexible and responsive to functionality needs.

Suggestion

For further development, it is recommended that the data pipeline be changed from batch processing (CSV) to real-time integration with log sources (e.g., Syslog Server or firewall data streams) to enable anomaly detection and alerting as soon as they occur. Furthermore, it is recommended to add an automated notification feature (e.g., via email or Telegram API) to send alerts to UPA-TIK Staff as soon as an activity with an Anomaly Score above the CRITICAL threshold is detected, as well as implementing a persistent database to store the Live Scan query history to build a local UNIMA threat database.

Acknowledgment

The author would like to express His gratitude to God Almighty for His blessings and grace so that this Research Report can be completed. The author would also like to thank Sondy C. Kumajas, S.T., M.T., as Academic Supervisor, and Quido C. Kainde, S.T., M.T., as Head of the Information and Communication Technology Academic Support Unit of Manado State University as well as Research Supervisor, for their full support and guidance. I would also like to express my gratitude to my mother and sister who have always been my pillars of prayer, encouragement, motivation and unwavering support throughout my research program.

References

Abibulaiev, A., Pukach, P., & Vovk, M. (2026). Context-Aware ML/NLP Pipeline for Real-Time Anomaly Detection and Risk Assessment in Cloud API Traffic. *Machine Learning and Knowledge Extraction*, *8*(1), 25. https://doi.org/10.3390/make8010025

Al-Akbar, M. 'Azam, Y. A. P., & Gurning, A. N. A. H. (2025). Deteksi trafik anomali berdasarkan pola trafik menggunakan isolation forest. *Cosmic Jurnal Teknik, 2*(3), 88–95.

Aminu, M., Akinsanya, A., Dako, D. A., & Oyedokun, O. (2024). Enhancing cyber threat detection through real-time threat intelligence and adaptive defense mechanisms. *International Journal of Computer Applications Technology and Research*, *13*(8), 11-27. https://doi.org/10.7753/IJCATR1308.1002

Artika, D. A., Rumahorbo, D., Al-Majid, M. H., & Kiswanto, D. (2025). Implementasi sistem keamanan website dengan analisis log dan deteksi aktivitas anomali menggunakan isolation forest. *Jurnal Informatika dan Teknik Elektro Terapan, 13*(3S1), 1868–1877. https://journal.eng.unila.ac.id/index.php/jitet/article/view/8133

Aryanti, N., Ardiansyah, W., & Anggaira, A. S. (2023). Student perceptions toward eight forms of independent learning activities in independent learning independent campus

ISSN: 2716-3865 (Print), 2721-1290 (Online)

program (mbkm). *International Journal of Research in Vocational Studies (IJRVOCAS)*, *3*(3), 52-62. https://doi.org/10.53893/ijrvocas.v3i3.233

Bainuan, L. D., & Tarigan, Y. Z. (2024). Strategy of the independent learning program-independent campus (MBKM) for the health study program curriculum: Scoping review. *Journal of Scientific Research, Education, and Technology (JSRET)*, *3*(1), 19-31. https://doi.org/10.58526/jsret.v3i1.310

Bianco, A. (2024). *Automatic Cybersecurity Risk Analysis* (Doctoral dissertation, Politecnico di Torino).

Das, S., & Nayak, T. (2013). Impact of cybercrime: Issues and challenges. *International journal of engineering sciences & Emerging technologies*, *6*(2), 142-153. https://doi.org/10.31142/ijtsrd23456

Dianita Pramesti, K., Meisya, N. I., & Amrillah, R. (2024). Relevansi lulusan perguruan tinggi dengan dunia kerja. *An Najah Jurnal Pendidikan Islam dan Sosial Agama, 3*(4), 236–243. https://journal.nabest.id/index.php/annajah

Ekundayo, F., Atoyebi, I., Soyele, A., & Ogunwobi, E. (2024). Predictive analytics for cyber threat intelligence in fintech using big data and machine learning. *Int J Res Publ Rev*, *5*(11), 1-15. https://doi.org/10.55248/gengpi.5.1124.3352

Fadli, M., Hanum, L., Amri, K., & Rusli, R. (2024). Barriers and Strategies: Analysis of the Implementation of Independent Learning Independent Campus (MBKM) at PTKI in Aceh. *QALAMUNA: Jurnal Pendidikan, Sosial, dan Agama*, *16*(2), 1101-1114. https://doi.org/10.37680/qalamuna.v16i2.5730

Ismanda, R. S., Silitonga, M. T. A., & Hasanah, S. N. (2025). Deteksi hybrid anomali transaksi digital dengan optimasi isolation forest-K-means untuk peningkatan keamanan finansial. *INNOVATIVE: Journal of Social Science Research, 5*(3). https://doi.org/10.31004/innovative.v5i3.19791

Issenoro, Trisnawati, H., Tarigan, S. O., & Faizah, N. M. (2025). Web-based network anomaly detection system for disaster recovery center: A SIEM implementation at the Indonesian Attorney General Training Agency. *Journal Innovations Computer Science, 4*(1), 1–17. Yayasan Kawanad. https://doi.org/10.56347/jics.v4i1.217

Kamalia, P. U., & Andriansyah, E. H. (2021). Independent learning-independent campus (MBKM) in studentsâ€™ perception. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di bidang Pendidikan, Pengajaran, dan Pembelajaran*, *7*(4), 857-867. https://doi.org/10.33394/jk.v7i4.4031

Kaul, D., & Khurana, R. (2021). AI to detect and mitigate security vulnerabilities in APIs: encryption, authentication, and anomaly detection in enterprise-level distributed systems. *Eigenpub Review of Science and Technology*, *5*(1), 34-62.

Kraemer-Mbula, E., Tang, P., & Rush, H. (2013). The cybercrime ecosystem: Online innovation in the shadows?. *Technological Forecasting and Social Change*, *80*(3), 541-555. https://doi.org/10.1016/j.techfore.2012.07.002

Lima, M., Viana, C., Santos, W. R., Neves, F., Campos, J. R., & Aires, F. (2025). Toward using cyber threat intelligence with machine and deep learning for IoT security: a comprehensive study. *The Journal of Supercomputing*, *81*(15), 1-39. https://doi.org/10.1007/s11227-025-07850-2

Mangkey, R. L. B., Rorimpandey, G. C., & Kumajas, S. C. (2025). Analisis prostitusi online pada aplikasi Michat menggunakan algoritma naïve Bayes dan framework NIST.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Jurnal Minfo Polgan, 14(2), 2658–2666. Politeknik Ganesha. https://doi.org/10.33395/jmp.v14i2.15466

Naseer, I. (2023). Machine learning applications in cyber threat intelligence: a comprehensive review. *The Asian Bulletin of Big Data Management*, *3*(2), 190-200. https://doi.org/10.62019/abbdm.v3i2.85

Putri, C. L. S., & Rachman, R. (2025). Deteksi anomali pembayaran TPD dan TKGB dengan isolation forest dan evaluasi risiko berbasis COSO ERM. *Jurnal Ilmiah Informatika Global, 16*(2), 307–312.

Rambing, G. J. J., Kumajas, S. C., & Santa, K. (2024). Penerapan business intelligence pada upa teknologi informasi dan komunikasi Universitas Negeri Manado. *Journal of Informatics, Business, Education, and Innovation Technology, 2*, 104–114.

Rehman, F., & Hashmi, S. (2023). Enhancing cloud security: A comprehensive framework for real-time detection analysis and cyber threat intelligence sharing. *Advances in Science, Technology and Engineering Systems Journal*, *8*(6), 107-119. https://doi.org/10.25046/aj080612

Sharma, G., Vidalis, S., Menon, C., Anand, N., & Kumar, S. (2021). Analysis and implementation of threat agents profiles in semi-automated manner for a network traffic in real-time information environment. *Electronics*, *10*(15), 1849. https://doi.org/10.3390/electronics10151849

Soumik, M. S., Mamun, K. S. A., Omim, S., Khan, H. A., & Sarkar, M. (2024). Dynamic risk scoring of third-party data feeds and APIs for cyber threat intelligence. *Journal of Computer Science and Technology Studies, 6*(1), 282–292. https://creativecommons.org/licenses/by/4.0/

Tambingon, H., & Tangkere, T. F. (2025). Optimizing Family Welfare Education Management through Information Technology: A Case Study Approach. *International Journal of Information Technology and Education*, *4*(2), 143-156.

Tjaija, A. (2022). Implementation of 'freedom to learn, independent campus'(MBKM) policy. *Al-Ishlah: Jurnal Pendidikan*, *14*(1), 319-328. https://doi.org/10.35445/alishlah.v14i1.2115

Vhalery, R., Setyastanto, A. M., & Leksono, A. W. (2022). Kurikulum Merdeka Belajar Kampus Merdeka: Sebuah kajian literatur. *Research and Development Journal of Education, 8*(1), 185. Universitas Indraprasta PGRI. http://dx.doi.org/10.30998/rdje.v8i1.11718

ISSN: 2716-3865 (Print), 2721-1290 (Online)