

LITERATURE REVIEW: PENERAPAN MACHINE LEARNING UNTUK DETEKSI INTRUSI PADA KEAMANAN WEB UMKM

Arjuna Putra Dwiyan Syah¹, Tining Haryanti², Luqman Hakim³

Universitas Muhammadiyah Surabaya

dwiyans26@gmail.com

ABSTRAK

Keamanan web pada Usaha Mikro, Kecil, dan Menengah (UMKM) menjadi perhatian utama di tengah meningkatnya serangan siber di Indonesia yang mencapai 14,9 juta kasus pada tahun 2025. Dengan keterbatasan anggaran dan sumber daya manusia, UMKM membutuhkan solusi deteksi intrusi yang terjangkau dan efisien. Penerapan *Machine Learning* terbukti efektif mengenali pola serangan secara otomatis, namun kajian spesifik untuk konteks web UMKM masih sangat terbatas. *Literature review* ini bertujuan menganalisis tren, metode, dan tantangan penerapan *Machine Learning* untuk deteksi intrusi pada web UMKM. Metode yang digunakan adalah *Systematic Literature Review* dengan artikel dari database Scopus dan SINTA periode 2020–2026. Hasil review menunjukkan algoritma Random Forest, SVM, dan Decision Tree paling banyak digunakan, di mana Random Forest mencapai akurasi hingga 99,99%. Meskipun demikian, sebagian besar penelitian masih diuji pada dataset publik dan belum diimplementasikan langsung pada lingkungan web UMKM. Selain itu, jenis serangan yang diteliti masih terbatas pada Brute Force dan DDoS. Kebaruan penelitian ini terletak pada tawaran arah pengembangan deteksi intrusi yang berfokus pada karakteristik web UMKM, serta memberikan kontribusi praktis sebagai referensi strategi keamanan siber yang adaptif.

Kata Kunci: Deteksi Intrusi, Keamanan Web, Machine Learning, Random Forest, UMKM

ABSTRACT

Web security for Micro, Small, and Medium Enterprises (MSMEs) has become a priority given the rising number of cyberattacks in Indonesia, which reached 14.9 million cases in 2025. Facing budget and manpower constraints, MSMEs require affordable and efficient intrusion detection solutions. Although machine learning has proven effective in automatically identifying attack patterns, specific research regarding the MSME web context remains very limited. This literature review aims to analyze the trends, methods, and challenges associated with implementing machine learning for intrusion detection on MSME websites. A systematic literature review methodology was employed, utilizing articles from the Scopus and SINTA databases published between 2020 and 2026. The review results indicate that Random Forest, SVM, and Decision Tree are the most frequently used algorithms, with Random Forest achieving accuracy levels of up to 99.99%. However, most studies rely on public datasets and have not yet been implemented directly within actual MSME web environments. Furthermore, the types of attacks studied are largely limited to Brute Force and DDoS. The originality of this research lies in its proposed development roadmap for intrusion detection that specifically addresses the characteristics of MSME websites, while also offering practical contributions as a reference for adaptive cybersecurity strategies.

Keywords: Intrusion Detection, Web Security, Machine Learning, Random Forest, MSME

Pendahuluan

Serangan siber terhadap website Usaha Mikro, Kecil, dan Menengah (UMKM) di Indonesia menunjukkan peningkatan yang signifikan dalam beberapa tahun terakhir. Berdasarkan laporan Kaspersky Security Network, sepanjang tahun 2025 tercatat 14.909.665 serangan siber berbasis web di Indonesia, setara dengan 40.848 upaya serangan per hari Pernita Hestin Untari, 2026 . Angka ini menempatkan Indonesia pada posisi rentan karena 22,4% pengguna internet menghadapi ancaman online. UMKM sebagai tulang punggung ekonomi nasional menjadi target empuk karena keterbatasan sumber daya yang dimiliki, baik dari segi anggaran maupun tenaga ahli di bidang teknologi informasi.

UMKM menghadapi tantangan unik dalam melindungi aset digital mereka. Berbeda dengan perusahaan besar yang memiliki tim keamanan siber khusus, UMKM umumnya beroperasi dengan keterbatasan anggaran dan minimnya tenaga ahli IT [1]. Survei terhadap 43 pelaku UMKM kerajinan di Kota Semarang menunjukkan bahwa 80% masih belum memahami konsep ketahanan siber, sementara sisanya hanya memahami namun belum menerapkannya secara komprehensif [1]. Kondisi ini semakin diperparah dengan peringkat keamanan siber Indonesia yang berada di posisi terakhir di antara negara-negara G20 pada tahun 2022 [2].

Saat ini, tersedia dua kategori utama solusi keamanan siber di pasar. Pertama, solusi komersial yang ditawarkan oleh vendor seperti CrowdStrike atau SentinelOne yang memiliki fitur lengkap namun biaya lisensinya terlalu tinggi bagi UMKM. Kedua, solusi open-source seperti Wazuh, Snort, atau OSSEC yang tersedia gratis tetapi membutuhkan keahlian teknis tinggi untuk instalasi, konfigurasi, dan integrasinya [3]. Akibatnya, banyak UMKM

memilih untuk tidak mengimplementasikan sistem keamanan sama sekali atau hanya mengandalkan solusi seadanya.

Dalam pengembangan sistem deteksi intrusi, penerapan machine learning (ML) dan deep learning (DL) telah menunjukkan hasil yang menjanjikan. [4] dalam systematic review terhadap 41 studi menemukan bahwa algoritma machine learning seperti Random Forest, Support Vector Machine (SVM), dan Decision Tree memiliki performa yang sangat baik dalam mendeteksi serangan yang sudah dikenal polanya melalui pembelajaran dari data historis. [5] mengembangkan model ensemble learning berbasis Random Forest yang mencapai akurasi 99,99% pada dataset NSL-KDD dan CIC-IDS2017, membuktikan efektivitas algoritma ini untuk deteksi intrusi. [6] juga menegaskan bahwa pendekatan feature selection dan ensemble learning efektif untuk membangun lightweight IDS pada perangkat dengan sumber daya terbatas.

Namun, sebagian besar penelitian deteksi intrusi berbasis machine learning masih diuji pada dataset publik seperti NSL-KDD, CICIDS2017, atau UNSW-NB15, bukan pada lingkungan web UMKM yang sesungguhnya. Selain itu, cakupan jenis serangan lain seperti Port Scanning, SQL Injection, dan Cross-Site Scripting (XSS) masih jarang diteliti dalam konteks web UMKM [3]. Padahal, kelima jenis serangan ini merupakan ancaman yang paling umum terjadi pada web UMKM.

Berdasarkan uraian di atas, literature review ini bertujuan untuk menganalisis tren, metode, dan tantangan penerapan machine learning untuk deteksi intrusi pada keamanan web UMKM, serta mengidentifikasi kesenjangan penelitian yang dapat menjadi

arah pengembangan sistem deteksi intrusi di masa mendatang

Materi dan Metode

1. Jenis Penelitian

Penelitian ini merupakan systematic literature review (SLR) yang bertujuan untuk mengidentifikasi, mengevaluasi, dan mensintesis penelitian-penelitian terdahulu tentang penerapan machine learning untuk deteksi intrusi pada keamanan web UMKM. Systematic literature review dipilih karena metode ini memungkinkan peneliti untuk melakukan pencarian literatur secara terstruktur, transparan, dan dapat direplikasi oleh peneliti lain. Pendekatan ini sejalan dengan penelitian [4] yang juga menggunakan metode SLR dengan panduan PRISMA 2020 untuk menganalisis 41 studi tentang machine learning dan deep learning untuk intrusion detection system.

2. Sampel dan Populasi

Populasi dalam penelitian ini adalah seluruh artikel ilmiah yang membahas tentang deteksi intrusi, machine learning, dan keamanan web yang dipublikasikan pada database terindeks Scopus dan SINTA periode 2020-2026. Sampel penelitian diambil dengan kriteria inklusi sebagai berikut: (1) artikel dipublikasikan pada periode 2020-2026; (2) artikel terindeks Scopus (Q1-Q4) atau SINTA (S1-S3); (3) artikel membahas topik deteksi intrusi atau intrusion detection system (IDS); (4) artikel menggunakan metode machine learning (Random Forest, Support Vector Machine, Decision Tree, atau ensemble learning); (5) artikel relevan dengan keamanan web atau web UMKM. Proses seleksi artikel dilakukan melalui beberapa tahap: identifikasi, screening, eligibility, dan inclusion. Berdasarkan proses tersebut, terpilih 33 artikel yang memenuhi kriteria inklusi untuk dianalisis lebih lanjut. Proses ini mengadaptasi

pendekatan yang digunakan oleh [4] dalam systematic review mereka.

Hasil dan Pembahasan

Hasil

Berdasarkan proses pencarian literatur pada database Scopus dan SINTA periode 2020-2026 dengan kata kunci deteksi intrusi, machine learning, dan keamanan web, ditemukan sebanyak 33 artikel yang memenuhi kriteria inklusi. Hasil review terhadap 33 artikel tersebut disajikan dalam tiga bagian: distribusi artikel berdasarkan tahun publikasi, distribusi berdasarkan metode yang digunakan, dan ringkasan temuan utama.

Distribusi Artikel Berdasarkan Tahun Publikasi

Sebaran tahun publikasi artikel yang dianalisis menunjukkan peningkatan jumlah penelitian tentang deteksi intrusi berbasis machine learning dalam lima tahun terakhir. Tabel 1 menyajikan distribusi artikel berdasarkan tahun publikasi.

Table 1. Distribusi Artikel Berdasarkan Tahun Publikasi

Tahun	Jumlah Artikel	(%)
2020		3,0
2021		6,1
2022		9,1
2023		15,2
2024		24,2
2025	14	42,4
Total	33	100

Berdasarkan Tabel 1, terjadi peningkatan jumlah publikasi yang signifikan pada tahun 2024 dan 2025, yang mengindikasikan bahwa topik deteksi intrusi berbasis machine learning semakin mendapat perhatian dari peneliti dalam dua tahun terakhir.

Distribusi Berdasarkan Metode Machine Learning

Dari 33 artikel yang dianalisis, algoritma machine learning yang paling banyak digunakan adalah Random Forest, Support Vector Machine (SVM), dan Decision Tree. Tabel 2 menyajikan distribusi frekuensi penggunaan algoritma machine learning.

Table 2. Distribusi Frekuensi Penggunaan Algoritma Machine Learning

Algoritma	Jumlah Artikel (%)	
Random Forest	8	24,2
Support Vector Machine (SVM)	6	18,2
Decision Tree	5	15,2
K-Nearest Neighbors (KNN)	4	12,1
LSTM / Deep Learning	4	12,1
XGBoost	3	9,1
Ensemble Learning (Stacking)	3	9,1
Total	33	100

Distribusi Berdasarkan Dataset yang Digunakan

Sebagian besar penelitian yang direview menggunakan dataset publik untuk menguji model deteksi intrusi yang dikembangkan. Tabel 3 menyajikan distribusi penggunaan dataset.

Table 3. Distribusi Penggunaan Dataset

Dataset	Jumlah Artikel(%)	
NSL-KDD	5	15,2
CIC-IDS2017	4	12,1
UNSW-NB15	5	9,1
Bot-IoT / TON_IoT	4	9,1
AWID2 / AWID3	4	6,1
KDDCup99	3	6,1
Dataset sendiri / Simulasi	3	9,1
Lainya	11	33,3
Total	33	100

Ringkasan Temuan Utama

Berdasarkan analisis terhadap 33 artikel, ditemukan beberapa temuan utama sebagai berikut. Pertama, algoritma Random Forest menunjukkan kinerja terbaik dengan akurasi mencapai 99,99% pada pengujian dengan dataset NSL-KDD dan CIC-IDS2017 [5]. Pendekatan ensemble learning dan feature selection terbukti efektif dalam meningkatkan akurasi sekaligus mengurangi konsumsi sumber daya komputasi.

Kedua, sebagian besar penelitian (90,9%) masih diuji pada dataset publik atau lingkungan laboratorium, bukan pada lingkungan web UMKM yang sesungguhnya. Hanya satu artikel [6] yang secara spesifik mengimplementasikan sistem deteksi intrusi pada konteks UMKM, itupun masih terbatas pada dua jenis serangan (Brute Force dan DDoS).

Ketiga, dari 33 artikel yang dianalisis, sebagian besar penelitian hanya berfokus pada deteksi serangan Brute Force dan DDoS. Serangan lain seperti Port Scanning, SQL Injection, dan Cross-Site Scripting (XSS) masih jarang diteliti dalam konteks web UMKM [3].

Keempat, penelitian yang mengintegrasikan deteksi, respons otomatis, dan notifikasi real-time dalam satu platform masih sangat terbatas. (2025) [3] adalah satu-satunya yang mengintegrasikan ketiganya, namun dengan cakupan serangan yang terbatas.

Pembahasan

Berdasarkan hasil review terhadap 33 artikel yang telah dianalisis, pembahasan difokuskan pada tiga aspek utama: (1) kinerja algoritma machine learning untuk deteksi intrusi, (2) kesenjangan implementasi pada lingkungan

web UMKM, dan (3) integrasi deteksi, respons otomatis, dan notifikasi *real-time*.

Kinerja Algoritma Machine Learning untuk Deteksi Intrusi

Dari 33 artikel yang dianalisis, algoritma Random Forest muncul sebagai metode yang paling banyak digunakan dan menunjukkan kinerja terbaik. Nassreddine et al. (2025) [5] dalam penelitiannya berhasil mencapai akurasi 99,99% pada pengujian dengan dataset NSL-KDD dan CIC-IDS2017 menggunakan pendekatan ensemble learning stacking yang menggabungkan Random Forest, XGBoost, dan Decision Tree dengan logistic regression sebagai meta-model. Keberhasilan ini tidak terlepas dari penggunaan teknik feature selection yang menggabungkan Correlation-based Feature Selection (CFS) dan embedded feature selection berbasis XGBoost, serta teknik resampling SMOTE-TOMEK untuk menangani ketidakseimbangan dataset.

Temuan ini sejalan dengan systematic review yang dilakukan oleh [4], yang menyatakan bahwa algoritma Random Forest memiliki keunggulan dalam menangani dataset berdimensi tinggi, ketahanan terhadap noise dan overfitting, serta interpretabilitas yang baik dibandingkan dengan algoritma deep learning yang cenderung menjadi "*black box*") [6] juga menegaskan bahwa pendekatan feature selection dan ensemble learning efektif untuk membangun lightweight IDS pada perangkat dengan sumber daya terbatas, yang relevan dengan karakteristik infrastruktur UMKM.

Sementara itu, algoritma Support Vector Machine (SVM) dan Decision Tree juga cukup populer, namun kinerjanya masih kalah dibandingkan Random Forest, terutama dalam menangani dataset yang besar dan tidak seimbang. Penelitian [7] yang menggunakan SVM untuk deteksi intrusi *real-time* hanya

mencapai akurasi 90,14%, jauh di bawah capaian Random Forest yang mencapai 99,99%. Hal ini menunjukkan bahwa pendekatan ensemble learning seperti Random Forest lebih unggul dibandingkan algoritma tunggal seperti SVM atau Decision Tree.

Kesenjangan Implementasi pada Lingkungan Web UMKM

Meskipun berbagai penelitian telah berhasil mengembangkan model deteksi intrusi dengan akurasi tinggi, sebagian besar penelitian tersebut masih diuji pada dataset publik seperti NSL-KDD, CIC-IDS2017, atau UNSW-NB15. Hanya satu artikel yang secara spesifik mengimplementasikan sistem deteksi intrusi pada konteks UMKM, yaitu penelitian yang dilakukan oleh [3]. Penelitian tersebut mengintegrasikan Wazuh sebagai SIEM, Fail2Ban untuk respons otomatis, Telegram API untuk notifikasi *real-time*, dan Random Forest untuk klasifikasi serangan, dengan metodologi pengembangan PPDIOO.

Namun, penelitian [3] masih memiliki keterbatasan pada cakupan jenis serangan yang diuji, yaitu hanya terbatas pada Brute Force dan DDoS. Padahal, berdasarkan data dan literatur, web UMKM juga rentan terhadap serangan lain seperti Port Scanning, SQL Injection, dan Cross-Site Scripting (XSS). Keterbatasan ini menjadi celah penelitian yang perlu diisi oleh penelitian selanjutnya.

Kondisi ini menunjukkan adanya kesenjangan antara penelitian yang dilakukan di lingkungan laboratorium atau dataset publik dengan implementasi di dunia nyata pada web UMKM. Web UMKM memiliki karakteristik lalu lintas yang berbeda, keterbatasan sumber daya komputasi, serta minimnya tenaga ahli IT yang menjadi tantangan tersendiri dalam implementasi sistem deteksi intrusi. Penelitian [1] mengidentifikasi bahwa 80% pelaku UMKM di Kota Semarang masih belum

memahami konsep ketahanan siber, sementara sisanya hanya memahami namun belum menerapkannya secara komprehensif.

Integrasi Deteksi, Respons Otomatis, dan Notifikasi Real-Time

Salah satu temuan penting dari literature review ini adalah masih terbatasnya penelitian yang mengintegrasikan ketiga komponen utama sistem keamanan: deteksi, respons otomatis, dan notifikasi real-time dalam satu platform terintegrasi. Penelitian [7] berhasil mengintegrasikan Snort sebagai IDS dengan WhatsApp API untuk notifikasi real-time, namun tidak dilengkapi dengan mekanisme respons otomatis. Sistem hanya bersifat monitoring dan memberikan peringatan ke administrator, tetapi tidak secara otomatis memblokir serangan yang terdeteksi.

Sebaliknya, penelitian [8] berhasil mengimplementasikan mekanisme respons otomatis untuk memblokir serangan (SYN Flood, backdoor, brute-force) menggunakan firewall filtering dan Intrusion Prevention System (IPS), namun tidak dilengkapi dengan notifikasi real-time. Penelitian [9] bahkan menggunakan metodologi PPDIOO yang sama dengan [6], menunjukkan bahwa metodologi ini dapat diterapkan untuk berbagai jenis sistem keamanan.

Penelitian (Ariyanto et al., 2025) menjadi satu-satunya yang berhasil mengintegrasikan ketiga komponen tersebut (deteksi, respons otomatis, dan notifikasi) sekaligus dengan metrik kinerja yang jelas. Sistem tersebut mencapai MTTD 4,7 detik untuk Brute Force dan 7,3 detik untuk DDoS, serta MTTR kurang dari 10 detik. Namun, integrasi ini masih terbatas pada cakupan serangan yang sempit, sehingga belum dapat melindungi web UMKM secara komprehensif.

Implikasi untuk Penelitian Mendatang

Berdasarkan pembahasan di atas, terdapat beberapa implikasi untuk penelitian mendatang. Pertama, perlu dikembangkan model deteksi intrusi berbasis machine learning yang diuji secara langsung pada lingkungan web UMKM, bukan hanya pada dataset publik. Kedua, cakupan jenis serangan perlu diperluas dari hanya Brute Force dan DDoS menjadi mencakup Port Scanning, SQL Injection, dan Cross-Site Scripting (XSS). Ketiga, sistem keamanan yang dikembangkan harus mengintegrasikan deteksi, respons otomatis, dan notifikasi real-time dalam satu platform yang mudah diimplementasikan oleh UMKM dengan keterbatasan sumber daya. Keempat, dokumentasi arsitektur perangkat lunak perlu disajikan secara sistematis untuk memudahkan replikasi dan pengembangan lebih lanjut.

Kesimpulan

Literature review ini telah menganalisis 33 artikel dari database Scopus dan SINTA periode 2020-2026 tentang penerapan machine learning untuk deteksi intrusi pada keamanan web UMKM. Berdasarkan hasil analisis, ditemukan bahwa algoritma Random Forest, Support Vector Machine (SVM), dan Decision Tree menjadi metode yang paling banyak digunakan, dengan Random Forest mencapai akurasi tertinggi hingga 99,99% pada pengujian dengan dataset publik. Pendekatan ensemble learning dan feature selection terbukti efektif dalam meningkatkan akurasi sekaligus mereduksi konsumsi sumber daya komputasi. Namun, sebagian besar penelitian masih diuji pada dataset publik atau lingkungan laboratorium, sementara implementasi pada lingkungan web UMKM yang sesungguhnya masih sangat terbatas. Selain itu, cakupan jenis serangan yang diuji masih terbatas pada Brute Force dan DDoS,

sementara serangan lain seperti Port Scanning, SQL Injection, dan Cross-Site Scripting (XSS) masih jarang diteliti dalam konteks web UMKM.

Penelitian yang mengintegrasikan deteksi, respons otomatis, dan notifikasi real-time dalam satu platform juga masih sangat terbatas. Kebaruan literature review ini adalah menawarkan pengembangan arah penelitian deteksi intrusi berbasis machine learning yang lebih berfokus pada implementasi di lingkungan web UMKM. Kontribusi praktis penelitian ini sebagai referensi bagi peneliti selanjutnya dalam menentukan strategi pengembangan sistem deteksi intrusi yang sesuai dengan karakteristik dan keterbatasan UMKM.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Panitia Seminar Nasional dan *Call for Paper* Universitas Muhammadiyah Muara Bungo atas kesempatan yang diberikan sehingga penulis bisa menyelesaikan artikel ini sampai publikasi. Penulis juga terimakasih kepada semua yang terlibat dalam penelitian ini sehingga sampai artikel ini diterbitkan.

Daftar Pustaka

- [1] Putranti, I. R., Amaliyah, A., & Windiani, R. (2020). Smartcity: Model ketahanan siber untuk usaha kecil dan menengah. *Jurnal Ketahanan Nasional*, 26(2), 222-241.
- [2] Vincha, C., & Satrio, J. (2024). Kemunculan ancaman siber teknologi 5G dan implikasinya terhadap ketahanan siber Indonesia. *Jurnal Ketahanan Nasional*, 30(2), 222-241.
- [3] Ariyanto, Y., Syaifudin, Y. W., Ratsanjani, M. H., Muladawila, A. R., Fatmawati, T., Saputra, P. Y., & Setiadi, C. (2025). Cyber threat detection and automated response using Wazuh and Telegram API. *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, 25(1), 173-188.
- [4] Zmaimita, H., Madani, A., & Zine-Dine, K. (2025). Machine and deep learning for intrusion detection: A PRISMA-guided systematic review of recent advances. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 11(1), 66-74.
- [5] Nassreddine, G., Nassereddine, M., & Al-Khatib, O. (2025). Ensemble learning for network intrusion detection based on correlation and embedded feature selection techniques. *Computers*, 14(3), 82.
- [6] Fatima, M., Rehman, O., Rahman, I. M. H., Ajmal, A., & Park, S. J. (2024). Towards ensemble feature selection for lightweight intrusion detection in resource-constrained IoT devices. *Future Internet*, 16(10), 368.
- [7] Simanjuntak, J., & Rizal, S. (2025). Network security monitoring system via Whatsapp using Snort on Ubuntu. *Jurnal Jaringan Komputer dan Keamanan*, 6(1), 9-21.
- [8] Pranata, V. (2025). Next-generation firewall design using an intrusion prevention system (IPS) method for securing the web portal server of Universitas Bina Darma. *Jurnal Jaringan Komputer dan Keamanan*, 6(2), 87-92.
- Mubarok, M. S. H. (2025). Implementation of a network security system against SYN flood attacks and unauthorized access blocking using firewall filtering in the IT service department of PT Pupuk Sriwidjaja Palembang. *Jurnal Jaringan Komputer dan Keamanan*, 6(2), 78-86.