JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 4, No. 3, November 2025, pp. 772 - 791
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/151

Literature Review: A Machine Learning Approach to Web Attack Detection

Nazli Rahmeisi*[1], Eksa umar Gani[2], Arief Arfriandi[3]
[1,2,3] Fakultas Teknik, Universitas Negeri Semarang
Jl. Sekaran, Gunungpati, 50229, Kota Semarang, Jawa Tengah, Indonesia
email: nazlirahmeisi@students.unnes.ac.id*[1]; eksaumar80@students.unnes.ac.id[2];
arfriandi@mail.unnes.ac.id[3]

ARTICLE INFO

Article History:

Received: 22-09-2025
Revised: 05-10-2025
Accepted: 06-11-2025
Published: 21-11-2025

ABSTRACT

The rapid growth of web technologies and online services has increased the exposure of web applications to cyber threats such as Cross-Site Scripting (XSS) and SQL Injection (SQLi). Conventional rule-based mechanisms, such as Web Application Firewalls (WAFs), often fail to detect emerging attack patterns. To address this, Machine Learning (ML) and Deep Learning (DL) have emerged as adaptive approaches for enhancing web attack detection. This study performs a Systematic Literature Review (SLR) following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to analyze recent ML/DL-based detection methods. Of the 263 retrieved studies, 15 met the inclusion criteria for detailed review. The findings reveal that Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) are the most applied algorithms. At the same time, recent works emphasize Transformer-based and hybrid ML–DL models. These approaches achieved robust performance (accuracy 85–97%, F1-score >90%) but still face challenges in dataset representativeness, class imbalance, and computational cost. This review highlights future research directions in Explainable Artificial Intelligence (XAI), Federated Learning (FL), and adversarial robustness to develop more efficient and trustworthy web attack detection systems.

Keywords: *Web attack detection, Machine Learning, Deep Learning, Intrusion Detection System (IDS), Transformer.*

Abstrak

Perkembangan pesat *web technologies* dan layanan daring meningkatkan kerentanan *web applications* terhadap *cyber threats* seperti *Cross-Site Scripting (XSS)* dan *SQL Injection (SQLi)*. Mekanisme keamanan konvensional berbasis aturan, seperti *Web Application Firewalls (WAFs)*, sering kali gagal mendeteksi pola serangan baru. Untuk mengatasi keterbatasan tersebut, pendekatan *Machine Learning (ML)* dan *Deep Learning (DL)* muncul sebagai metode adaptif dalam meningkatkan efektivitas *web attack detection*. Penelitian ini melakukan *Systematic Literature Review (SLR)* berdasarkan pedoman *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* 2020 guna menganalisis metode deteksi berbasis ML/DL terkini. Dari 263 artikel yang diidentifikasi, 15 artikel memenuhi kriteria untuk dianalisis secara mendalam. Hasil menunjukkan bahwa algoritma *Random Forest (RF)*, *Support Vector Machine (SVM)*, *Convolutional Neural Network (CNN)*, dan *Long Short-Term Memory (LSTM)* paling sering digunakan,

DOI: https://doi.org/10.51903/3w0vwc80

p-ISSN : 2803-1507 e-ISSN : 2803-1531

dengan tren riset mengarah pada model *Transformer* dan hibrid ML–DL. Model tersebut menunjukkan performa kuat (akurasi 85–97%, *F1-score* >90%), meskipun masih menghadapi tantangan seperti representativitas *dataset*, ketidakseimbangan kelas, dan biaya komputasi. Studi ini menyoroti arah riset masa depan pada *Explainable Artificial Intelligence (XAI)*, *Federated Learning (FL)*, dan *adversarial robustness* untuk membangun sistem deteksi serangan web yang lebih efisien dan andal.

Kata Kunci: *Deteksi Serangan Web, Machine Learning, Deep Learning, Intrusion Detection System (IDS), Transformer.*

1. PENDAHULUAN

Perkembangan pesat teknologi web dan peningkatan jumlah layanan daring (*online services*) telah secara signifikan memperluas permukaan serangan siber. Aplikasi web modern, yang menjadi tulang punggung berbagai layanan *e-commerce*, perbankan digital, hingga *e-government*, semakin rentan terhadap ancaman keamanan yang kompleks dan canggih. Jenis serangan seperti *Web Shell*, *Cross-Site Scripting (XSS)*, *SQL Injection (SQLi)*, serta manipulasi *HTTP request* kini semakin bersifat otomatis dan adaptif, sehingga menimbulkan tantangan serius bagi pendekatan keamanan tradisional. Mekanisme berbasis aturan (*rule-based firewall*) maupun *Web Application Firewall (WAF)* konvensional sering kali tidak mampu mendeteksi varian baru serangan dengan cepat dan akurat, terutama ketika pola serangan terus berevolusi [1], [2].

Secara global, tren ancaman siber menunjukkan eskalasi yang signifikan. Laporan *ENISA Threat Landscape 2023* menegaskan bahwa serangan terhadap aplikasi web, termasuk *SQLi* dan *XSS*, masih mendominasi kategori insiden kritis dengan peningkatan serangan berbasis *exploitation* sebesar 25% dibandingkan tahun sebelumnya. *Cybersecurity Ventures* bahkan memproyeksikan kerugian ekonomi akibat kejahatan siber akan mencapai lebih dari USD 10,5 triliun per tahun pada 2025, menjadikannya salah satu ancaman terbesar bagi perekonomian digital global. Sementara itu, *Verizon Data Breach Investigations Report (DBIR)* 2024 mencatat bahwa lebih dari 40% dari seluruh pelanggaran data melibatkan eksploitasi kerentanan pada aplikasi web, terutama melalui serangan injeksi dan *credential stuffing*. Fakta-fakta ini menegaskan bahwa keamanan aplikasi web bukan lagi sekadar isu teknis, tetapi juga tantangan strategis bagi keberlanjutan bisnis dan layanan publik.

Selain itu, munculnya *Internet of Things (IoT)* dan *Software Defined Networks (SDN)* semakin memperluas vektor serangan. Beragam perangkat *IoT* yang memiliki keterbatasan sumber daya tetap dituntut untuk terlindungi, sementara jaringan *SDN* menghadirkan fleksibilitas sekaligus kompleksitas baru yang dapat dieksploitasi oleh penyerang [3], [4], [5]. Dalam konteks ini, riset terkini banyak berfokus pada pemanfaatan *Machine Learning (ML)* dan *Deep Learning (DL)* sebagai pendekatan alternatif yang lebih adaptif [6], [7], [8]. Beberapa studi menunjukkan hasil menjanjikan, misalnya [4], yang melaporkan model *DL* dengan tingkat *accuracy* tinggi dalam mendeteksi intrusi pada *IoT/SDN*, serta [3], yang memanfaatkan model hibrid *ML–DL* untuk deteksi multi-lapisan dengan hasil yang signifikan.

Pendekatan lain yang mulai berkembang adalah *semi-supervised learning*, yang berusaha mengurangi ketergantungan pada data berlabel penuh. [4] mengusulkan metode berbasis *sequence reconstruction* untuk menganalisis trafik *HTTP* normal sehingga anomali dapat dikenali tanpa data anotasi yang ekstensif. Di sisi lain, penelitian [9] menekankan pentingnya teknik *feature selection* dan *oversampling* untuk mengatasi *class imbalance* yang kerap ditemui pada sistem deteksi intrusi berbasis *ML*. Hal ini menunjukkan bahwa isu kualitas data sama pentingnya dengan pemilihan algoritma.

Namun, meskipun banyak kemajuan telah dicapai, sejumlah tantangan masih membatasi efektivitas implementasi *ML* dan *DL* dalam deteksi serangan web. Pertama, sebagian besar penelitian masih mengandalkan *dataset* publik seperti *CICIDS2017* dan *UNSW-NB15* yang bersifat generik dan kurang merepresentasikan pola serangan web terbaru [10], [11]. Kedua, masalah *false positive* yang tinggi dan keterbatasan model untuk melakukan generalisasi lintas domain masih menjadi penghalang utama dalam adopsi praktis. Ketiga, kompleksitas komputasi model *DL*, khususnya berbasis *transformer*, menyulitkan penerapan *real-time detection* pada perangkat dengan sumber daya terbatas [2], [12].

Dengan latar belakang tersebut, penelitian ini berfokus pada kajian sistematis literatur terkini mengenai penerapan *ML* dan *DL* dalam deteksi serangan web. Tujuan penelitian adalah: (1) mengidentifikasi algoritma *ML/DL* yang paling banyak digunakan; (2) mengevaluasi kinerja model berdasarkan metrik umum seperti *accuracy*, *precision*, *recall*, dan *F1-score*; serta (3) mengidentifikasi tantangan utama yang masih dihadapi

p-ISSN : 2803-1507 e-ISSN : 2803-1531

dalam implementasi praktis. Kontribusi utama penelitian ini adalah memberikan pemetaan komprehensif terhadap tren terbaru, mengelaborasi *research gap*, serta menyajikan rekomendasi praktis yang dapat mendukung pengembangan sistem deteksi serangan web yang lebih efektif, efisien, dan adaptif terhadap dinamika ancaman siber di masa depan.

2. TINJAUAN PUSTAKA

2.1. *Machine Learning (ML) dalam Deteksi Serangan Web*

Pendekatan *Machine Learning (ML)* klasik seperti *Random Forest (RF)*, *Support Vector Machine (SVM)*, *Naïve Bayes*, dan *K-Nearest Neighbor (KNN)* telah banyak digunakan dalam deteksi serangan web, di mana *SVM* terbukti mencapai *accuracy* tinggi pada *SQL Injection (SQLi)* dan *Cross-Site Scripting (XSS)* [9], [13]. Sementara itu, *RF* menunjukkan performa yang lebih baik dibanding *KNN* dalam mengatasi masalah *overfitting* [10]. Meskipun efektif untuk pola serangan sederhana, algoritma *ML* tradisional cenderung terbatas dalam mendeteksi serangan kompleks yang berbasis *payload* teks dari *HTTP request* [14]. Pada banyak kasus, algoritma ini membutuhkan data berlabel yang banyak untuk mencapai hasil yang optimal, yang dapat menjadi kendala jika data tersebut sulit didapatkan.

2.2. *Deep Learning (DL) dan Model Hybrid untuk Keamanan Web*

Peningkatan kompleksitas serangan mendorong pemanfaatan *Deep Learning (DL)*, di mana *Convolutional Neural Network (CNN)* terbukti efektif dalam mendeteksi *XSS* dan *SQLi* dengan performa lebih baik dibanding metode *ML* klasik [11]. Kombinasi *CNN–Long Short-Term Memory (LSTM)* dilaporkan mampu meningkatkan *F1-score* hingga 95% pada *dataset UNSW-NB15* [3], sedangkan model berbasis *transformer* seperti *Bidirectional Encoder Representations from Transformers (BERT)* semakin populer untuk analisis *payload* berbahaya [2], [13]. Dengan demikian, riset menunjukkan adanya pergeseran tren dari *ML* klasik menuju *DL* dan model *hybrid* yang lebih adaptif terhadap serangan web modern. Selain itu, model *hybrid* yang menggabungkan CNN dan LSTM mampu meningkatkan deteksi serangan berjangka panjang dengan lebih baik, mengingat karakteristik serangan yang semakin beragam dan dinamis.

2.3. *Dataset dan Benchmark yang Digunakan dalam Penelitian*

Dataset berperan krusial dalam menentukan kinerja model deteksi. *CICIDS2017*, *UNSW-NB15*, dan *CSIC 2010 HTTP Dataset* masih menjadi rujukan utama dalam berbagai studi [9]. Namun, penelitian terbaru menyoroti keterbatasan *dataset* tersebut dalam merepresentasikan serangan kontemporer [4], [11]. Tantangan seperti *class imbalance* sering menurunkan performa model [5], sehingga sejumlah penelitian mengusulkan pembangunan *dataset* baru atau penerapan teknik *data augmentation* seperti *Synthetic Minority Oversampling Technique (SMOTE)* dan *Generative Adversarial Networks (GAN)* untuk memperbaiki distribusi data [13].

2.4. *Tren Explainable Artificial Intelligence (XAI) dalam Sistem Deteksi*

Belakangan ini, *Explainable Artificial Intelligence (XAI)* menjadi perhatian penting karena memungkinkan interpretasi keputusan model yang kompleks, terutama dalam *Intrusion Detection Systems (IDS)*. Teknik seperti *SHapley Additive exPlanations (SHAP)* dan *Local Interpretable Model-agnostic Explanations (LIME)* digunakan untuk memberikan transparansi hasil prediksi, sehingga meningkatkan kepercayaan praktisi keamanan terhadap model [15], [16]. Pendekatan *XAI* juga membantu mengidentifikasi fitur kritis dalam trafik jaringan, yang relevan untuk audit dan penerapan di dunia nyata (*real-world deployment*). Dalam prakteknya, XAI digunakan untuk mengurangi kompleksitas model yang sangat besar, sehingga hasil deteksi lebih mudah dipahami oleh para praktisi di lapangan.

2.5. *Federated Learning (FL) dan Edge AI untuk Deteksi Real-Time*

Federated Learning (FL) menawarkan pendekatan terdistribusi dalam pelatihan model tanpa perlu menggabungkan data mentah, sehingga menjaga privasi pengguna dan cocok diterapkan pada lingkungan *Internet of Things (IoT)*. Integrasi *FL* dengan *Edge Artificial Intelligence (Edge AI)* memungkinkan deteksi serangan dilakukan secara *real-time* di perangkat lokal, mengurangi *latency* sekaligus mempertahankan *accuracy* [17], [18]. Meskipun menjanjikan, beberapa tantangan masih muncul, termasuk heterogenitas data, *communication overhead*, serta kebutuhan optimisasi model agar tetap ringan untuk perangkat dengan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

sumber daya terbatas. Selain itu, FL memungkinkan penggunaan data secara efisien, tanpa perlu berbagi data mentah antar perangkat, yang meningkatkan keamanan data pribadi pengguna.

2.6. *Unsupervised Learning dalam Deteksi Serangan Baru*

Untuk menghadapi serangan *zero-day*, pendekatan *Unsupervised Learning* seperti *autoencoder*, *k-means clustering*, dan *isolation forest* semakin banyak diterapkan. Pendekatan ini tidak memerlukan data berlabel penuh, melainkan mempelajari pola normal dan mengklasifikasikan anomali sebagai serangan potensial [19], [20]. Penelitian terkini juga mengeksplorasi kombinasi *Unsupervised Learning* dengan *Meta-Learning* agar model dapat beradaptasi lintas domain, meskipun tantangan seperti tingginya *false positive* dan keterbatasan interpretabilitas masih menjadi hambatan implementasi. Metode ini memungkinkan pengenalan pola serangan yang sebelumnya tidak terlihat, meskipun proses pelatihan yang lebih lama dapat menjadi salah satu kekurangan.

2.7. *Tantangan dan Kesenjangan Penelitian (Research Gap)*

Meskipun penelitian *ML* dan *DL* untuk deteksi serangan web terus berkembang, masih terdapat sejumlah tantangan seperti keterbatasan *dataset* representatif [5], [21], [22], tingginya *false positive rate* [14], serta kebutuhan komputasi besar pada model berbasis *transformer* yang menghambat penerapan *real-time detection* [2]. Selain itu, riset terkait model ringan (*lightweight models*), *XAI*, serta integrasi *FL* dengan *Unsupervised Learning* masih sangat terbatas. Oleh karena itu, penelitian ini mengusulkan kajian sistematis menggunakan pendekatan *Systematic Literature Review (SLR)* berbasis *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* guna menyajikan peta komprehensif tentang tren, performa, dan keterbatasan riset terkini. Sebagian besar penelitian masih berfokus pada metode yang hanya dapat diterapkan pada dataset terbatas, yang mengarah pada keterbatasan dalam generalisasi terhadap serangan yang lebih bervariasi.

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan *Systematic Literature Review (SLR)* yang disusun berdasarkan pedoman *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* 2020. Pendekatan ini dipilih untuk menjamin bahwa proses identifikasi, seleksi, dan evaluasi literatur dilakukan secara sistematis, transparan, serta dapat direplikasi. Dengan metode ini, analisis penerapan algoritma *Machine Learning (ML)* dan *Deep Learning (DL)* dalam deteksi serangan web dilakukan secara terstruktur sehingga menghasilkan temuan yang valid dan memberikan kontribusi ilmiah yang relevan bagi pengembangan *cybersecurity* berbasis web. Pendekatan ini juga memungkinkan peneliti untuk memetakan tren penelitian terkini dan mengidentifikasi kesenjangan yang belum banyak dikaji dalam topik deteksi serangan web.

3.1. *Research Question Definition*

Pertanyaan riset ini dirancang agar hasil kajian dapat memberikan pemahaman mendalam mengenai algoritma, performa, dan tantangan implementasi pada deteksi serangan web. Rumusan ini membantu memperjelas arah penelitian dan memastikan bahwa proses peninjauan literatur dilakukan secara fokus. Setiap pertanyaan berfungsi sebagai dasar dalam proses pengumpulan dan analisis data literatur yang relevan. Penelitian ini difokuskan untuk menjawab beberapa pertanyaan riset (*Research Questions*/RQ) berikut:

1. **RQ1:** Algoritma *Machine Learning (ML)* dan *Deep Learning (DL)* apa saja yang paling banyak digunakan dalam deteksi serangan web?
2. **RQ2:** Bagaimana performa algoritma tersebut dalam mendeteksi serangan (misalnya berdasarkan *accuracy*, *recall*, *precision*, *F1-score*, atau metrik lainnya)?
3. **RQ3:** Tantangan apa saja yang sering muncul dalam penerapan *ML/DL* untuk deteksi serangan web (misalnya keterbatasan *dataset*, *class imbalance*, efisiensi komputasi, atau *false positive*)?

3.2. *PICOC Framework*

Kerangka *Population, Intervention, Comparison, Outcome, Context (PICOC)* digunakan untuk merumuskan ruang lingkup kajian. Pendekatan *PICOC* membantu peneliti dalam mengidentifikasi elemen-elemen utama yang relevan dengan penelitian, seperti populasi kajian, jenis intervensi, dan hasil yang diharapkan. Dengan kerangka ini, batasan penelitian dapat ditentukan secara eksplisit sehingga proses seleksi literatur lebih terarah. Setiap elemen dalam kerangka ini diuraikan secara sistematis agar hubungan antar-komponen penelitian dapat dipahami dengan jelas. Rincian lengkap mengenai elemen *PICOC* yang digunakan dalam

p-ISSN : 2803-1507 e-ISSN : 2803-1531

penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. PICOC Framework

Elemen	Deskripsi
Population	Artikel penelitian yang membahas deteksi serangan web menggunakan *ML/DL*.
Intervention	Algoritma *ML* (misalnya *Random Forest (RF)*, *Support Vector Machine (SVM)*, *XGBoost*) dan *DL* (misalnya *Convolutional Neural Network (CNN)*, *Long Short-Term Memory (LSTM)*, *Transformer*, atau *hybrid*).
Comparison	Perbandingan dengan metode tradisional seperti *rule-based*, *signature-based*, atau *ensemble learning*.
Outcome	Kinerja model (*accuracy*, *precision*, *recall*, *F1-score*, *AUC*) dan tantangan implementasi.
Context	Studi yang terindeks di *Scopus*, *SINTA*, atau *Web of Science (WoS)* dan diterbitkan pada periode 2020–2025.

3.3. *Eligibility Criteria*

Untuk menjamin kualitas dan relevansi literatur, penelitian ini menetapkan kriteria kelayakan sebagai berikut: publikasi akademik terkini (2020–2025) yang membahas penerapan *ML* dan *DL* dalam *web attack detection*, ditulis dalam bahasa Inggris atau Indonesia, tersedia dalam *full-text*, serta dipublikasikan pada jurnal atau prosiding terindeks. Kriteria ini disusun untuk memastikan bahwa setiap artikel yang dikaji memiliki kontribusi ilmiah yang signifikan dan relevan dengan konteks penelitian. Pemilihan periode waktu yang terbatas juga dimaksudkan agar hasil kajian mencerminkan perkembangan teknologi terbaru dalam bidang keamanan web. Selain itu, hanya publikasi yang dapat diakses secara penuh yang dipertimbangkan agar analisis dapat dilakukan secara komprehensif. Pendekatan ini memastikan bahwa literatur yang digunakan benar-benar representatif terhadap konteks *cybersecurity* modern.

Proses *exclusion* dilakukan secara bertahap, mulai dari penghapusan artikel tanpa *full-text*, penyaringan terhadap studi sekunder seperti *review* atau *survey*, hingga penilaian kualitatif berdasarkan relevansi tematik dan ketelitian metodologis (evaluasi judul, *abstract*, dan isi *full-text*). Hanya artikel yang secara langsung menjawab *research question (RQ)* yang dipertahankan agar kajian mampu memberikan gambaran komprehensif dan mutakhir terkait perkembangan penelitian pada bidang ini. Tahapan seleksi ini dilakukan untuk memastikan hanya studi dengan kontribusi empiris dan metodologi yang kuat yang dimasukkan dalam proses analisis. Evaluasi menyeluruh juga dilakukan untuk menilai kejelasan tujuan penelitian dari setiap artikel. Kriteria inklusi dan eksklusi yang diterapkan secara rinci disajikan pada Tabel 2.

Tabel 2. Kriteria Inklusi dan Eksklusi

Kriteria	Inklusi	Eksklusi
Tahun publikasi	2020–2025	Sebelum 2020
Bahasa	Bahasa Inggris atau Indonesia	Bahasa selain Inggris/Indonesia
Jenis publikasi	Jurnal/prosiding terindeks *(SINTA, Scopus, WoS)*	Artikel non-akademik, *survey* sekunder
Aksesibilitas	*Full-text* tersedia	Hanya abstrak
Relevansi	Fokus pada ML/DL untuk deteksi serangan web	Topik selain keamanan web (misalnya *IDS* umum tanpa fokus web)

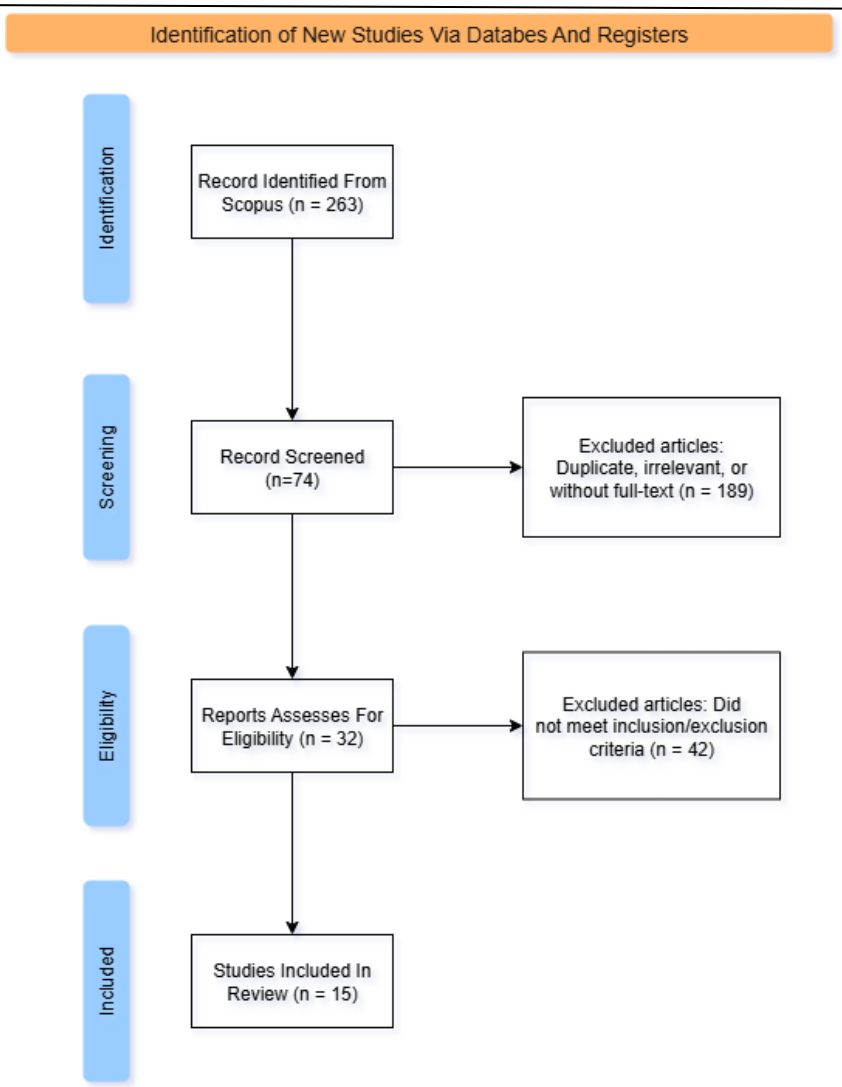
3.4. *Search Strategy and Selection Process*

Pencarian literatur dilakukan pada beberapa basis data utama, yaitu *Scopus*, *IEEE Xplore*, *SpringerLink*, *MDPI*, dan *Semantic Scholar*. Kata kunci yang digunakan merupakan kombinasi logika *Boolean* sebagai berikut: *"web attack detection" OR "web-based intrusion detection" OR "web security"* AND *"machine learning" OR "deep learning" OR "artificial intelligence"*. Selanjutnya, artikel-artikel yang ditemukan diseleksi melalui beberapa tahapan untuk menentukan artikel yang layak dianalisis. Proses seleksi dilakukan melalui tahapan sebagai berikut:

1. Tahap identifikasi: 263 artikel ditemukan dari lima basis data.
2. Tahap screening: 189 artikel dieliminasi karena duplikat, tidak relevan, atau tidak tersedia *full-text*.
3. Tahap eligibility: 32 artikel dievaluasi kelayakannya berdasarkan kriteria inklusi/eksklusi.
4. Tahap included: 15 artikel dipilih untuk dianalisis mendalam.

Pencarian literatur dilakukan secara sistematis pada beberapa basis data utama untuk memastikan cakupan artikel yang relevan. Setiap artikel yang memenuhi kriteria inklusi dievaluasi secara bertahap untuk mengeliminasi duplikat dan artikel yang tidak relevan. Proses ini memberikan dasar yang kokoh untuk tahap analisis lebih lanjut. Proses seleksi divisualisasikan dalam diagram alir PRISMA pada Gambar 1.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 1. Diagram alir PRISMA

3.5. *Data Extraction and Synthesis*

Tahap ekstraksi data dilakukan dengan mengumpulkan informasi spesifik yang berkaitan langsung dengan pertanyaan penelitian. Setiap data yang diperoleh dikelompokkan berdasarkan variabel yang telah ditentukan sebelumnya agar analisis lebih mudah dilakukan. Proses ini memastikan bahwa semua data penting dari setiap artikel tercatat dengan akurat sebelum dilakukan analisis lanjutan. Data diekstraksi secara sistematis dari 6 artikel terpilih, meliputi:

1. Algoritma ML/DL yang digunakan.
2. Jenis serangan web yang ditangani (misalnya *XSS, SQLi, phishing, ransomware*).
3. Dataset yang digunakan dalam evaluasi.
4. Metrik performa seperti *accuracy, precision, recall, F1-score, false positive rate.*

Hasil ekstraksi dianalisis secara kualitatif dan kuantitatif untuk menjawab setiap RQ. Analisis dilakukan dengan menyoroti tren algoritma, perbandingan performa, serta tantangan implementasi. Pendekatan sintesis ini memungkinkan identifikasi pola umum serta celah penelitian (*research gap*) di bidang deteksi serangan web berbasis ML/DL. Pendekatan ini digunakan untuk menyusun informasi dari berbagai artikel secara sistematis, sehingga tren penelitian dan tantangan implementasi dapat dipahami dengan lebih jelas.

3.6. *Data Extraction and Coding*

Dari 15 artikel terpilih, dilakukan proses ekstraksi data untuk menilai berbagai aspek penting dari penelitian

p-ISSN : 2803-1507 e-ISSN : 2803-1531

terkait. Setiap parameter dievaluasi secara mendetail untuk memastikan semua informasi penting terkait algoritma, jenis serangan web, dataset, metrik, dan tantangan implementasi terekam secara konsisten. Proses ekstraksi dilakukan berdasarkan parameter berikut:

1. Algoritma *ML/DL* yang digunakan.
2. Jenis serangan web yang ditangani (*SQLi*, *XSS*, *phishing*, *ransomware*).
3. *Dataset* yang digunakan (*CICIDS2017*, *UNSW-NB15*, *CSIC2010*, atau *custom dataset*).
4. Metrik performa (*accuracy*, *precision*, *recall*, *F1-score*, *AUC*, *false positive rate*).
5. Tantangan implementasi seperti *class imbalance*, kebutuhan komputasi tinggi, dan generalisasi lintas domain.

Untuk konsistensi analisis, dibuat kategori coding. Kategori ini dirancang untuk menyusun dan mengelompokkan data dari artikel yang dianalisis sehingga semua aspek penelitian dapat diklasifikasikan secara sistematis. Setiap sub-kategori dipilih berdasarkan relevansi terhadap algoritma, jenis serangan, dataset, metrik, dan tantangan implementasi. Pendekatan ini memudahkan proses *coding* sehingga analisis dapat dilakukan secara konsisten dan hasilnya dapat dibandingkan antarartikel. Kategori coding ditunjukkan pada Tabel 3.

Tabel 3. Coding Categories

Kategori	Sub-Kategori Utama
Algoritma	*RF, SVM, XGBoost, CNN, LSTM, Transformer, Hybrid, Ensemble*
Jenis serangan	*SQLi, XSS, phishing, ransomware, anomaly detection*
Dataset	*CICIDS2017, UNSW-NB15, CSIC2010, custom datasets*
Metrik	*Accuracy, Precision, Recall, F1-score, AUC, False Positive Rate*
Tantangan	*Class imbalance, dataset representatif, komputasi tinggi, generalisasi lintas domain*

3.7. *Data Synthesis*

Analisis dilakukan dengan tiga pendekatan yang saling melengkapi. Setiap pendekatan memberikan sudut pandang yang berbeda namun saling mendukung dalam memperoleh pemahaman menyeluruh terhadap hasil kajian. Ketiga pendekatan ini dipilih karena mampu melengkapi satu sama lain dalam mengungkap pola, efektivitas, serta tantangan dari hasil kajian literatur. Rincian pendekatan yang digunakan disajikan sebagai berikut:

1. Analisis kualitatif: mengidentifikasi tren algoritma, jenis serangan yang paling sering diteliti, dan *dataset* yang digunakan.
2. Analisis kuantitatif: menghitung distribusi artikel berdasarkan tahun, negara, dan metrik performa utama.
3. Thematic coding: mengelompokkan artikel sesuai RQ (RQ1, RQ2, RQ3) sehingga dapat memetakan tren, efektivitas, dan tantangan penelitian di bidang deteksi serangan web berbasis *ML/DL*.

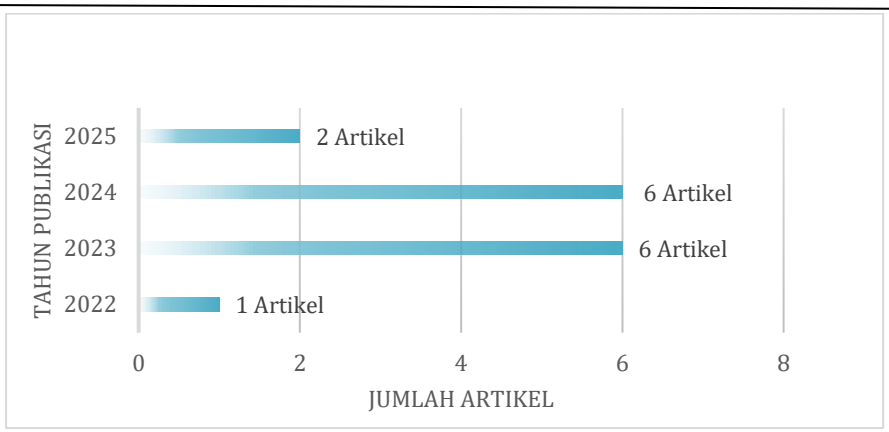
4. HASIL DAN PEMBAHASAN

Mengacu alur *Systematic Literature Review (SLR)* berdasarkan pedoman *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020*. Pada tahap *identification*, sebanyak 263 artikel diperoleh dari lima basis data utama, yaitu Scopus, IEEE Xplore, SpringerOpen, MDPI, dan Semantic Scholar, dengan menggunakan kombinasi kata kunci: *"Web Attack Detection"*, *"Machine Learning"*, *"Deep Learning"*, *"Intrusion Detection System (IDS)"*, dan *"SQL Injection/XSS Detection"*. Setelah tahap *screening* dilakukan berdasarkan judul dan abstrak, sebanyak 189 artikel dieliminasi karena duplikasi atau tidak relevan terhadap topik penelitian. Selanjutnya, pada tahap *eligibility*, 32 artikel dievaluasi secara mendalam untuk memastikan kesesuaian dengan kriteria inklusi, seperti periode publikasi (2020–2025), ketersediaan *full-text*, serta relevansi terhadap tema deteksi serangan web berbasis ML/DL. Akhirnya, 15 artikel utama dipilih untuk dianalisis secara tematik dan kuantitatif. Proses ini memastikan bahwa literatur yang dianalisis memiliki kualitas, keterbaruan, dan relevansi ilmiah yang tinggi terhadap tujuan penelitian.

4.1. *Distribusi Artikel*

Perkembangan penelitian tentang *Machine Learning* (*ML*) dan *Deep Learning* (*DL*) untuk deteksi serangan web menunjukkan peningkatan signifikan dari tahun ke tahun. Hal ini menunjukkan tingginya perhatian akademisi terhadap isu keamanan siber. Tren ini juga mencerminkan kebutuhan dunia nyata akan solusi yang lebih efektif dalam menghadapi ancaman siber global. Visualisasi tren publikasi tahunan ini dapat dilihat pada Gambar 2.

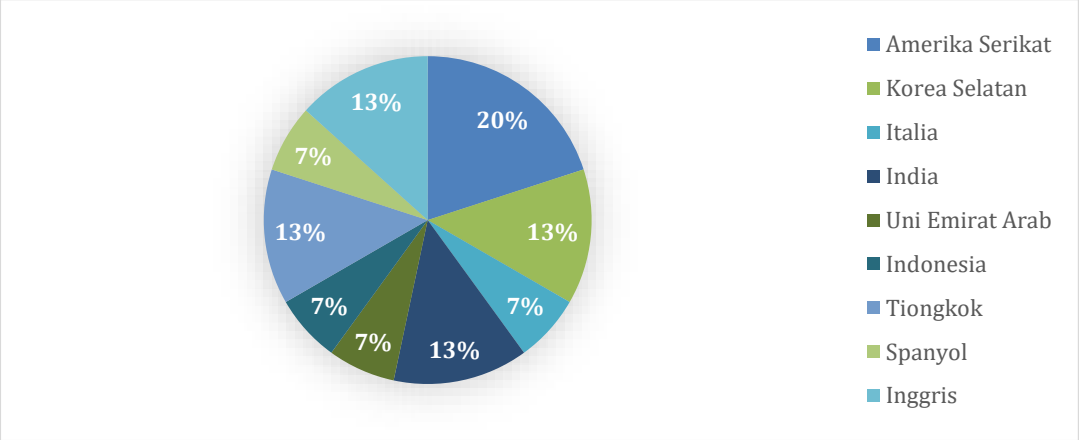
p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 2. Distribusi Artikel Berdasarkan Tahun Publikasi

Gambar 2 menunjukkan bahwa jumlah publikasi terbanyak terjadi pada tahun 2023 (6 artikel) dan 2024 (6 artikel), diikuti tahun 2025 (2 artikel) dan 2022 (1 artikel). Tren ini menunjukkan meningkatnya perhatian terhadap penerapan *Machine Learning (ML)* dan *Deep Learning (DL)* untuk deteksi serangan web, sejalan dengan eskalasi ancaman siber global yang mendorong intensifikasi riset di bidang *cybersecurity*. Pola peningkatan ini juga menunjukkan bahwa topik deteksi serangan web semakin mendapatkan dukungan dari komunitas riset internasional. Perkembangan metode dan dataset baru turut memperkaya arah penelitian di bidang ini.

Riset keamanan web berbasis *Artificial Intelligence* (*AI*) memiliki karakter internasional, melibatkan penulis dari berbagai negara. Kerja sama lintas negara ini memperkaya perspektif dan metode penelitian yang diterapkan. Kontribusi terbesar berasal dari Amerika Serikat, namun banyak negara lain juga aktif dalam bidang ini. Persebaran artikel menurut negara penulis dapat dilihat pada Gambar 3.



Gambar 3. Distribusi Artikel Berdasarkan Negara Penulis

Gambar 3 memperlihatkan distribusi artikel berdasarkan asal negara penulis. Amerika Serikat berkontribusi paling banyak (3 artikel), diikuti Korea Selatan, India, Italia, Uni Emirat Arab, dan Indonesia (masing-masing 1–2 artikel). Hal ini menegaskan bahwa riset keamanan web berbasis *Artificial Intelligence (AI)* bersifat global dan kolaboratif, melibatkan kerja sama lintas negara dan institusi seperti universitas, lembaga penelitian, serta industri teknologi. Keberagaman geografis ini juga berdampak pada variasi pendekatan metodologis yang digunakan dalam setiap penelitian. Faktor perbedaan infrastruktur digital dan prioritas keamanan di tiap negara turut memengaruhi fokus riset yang dilakukan.

Tabel 4 merangkum 15 artikel yang terpilih, meliputi judul, penulis/tahun, hasil penelitian, serta dataset/*tools* yang digunakan. Dari tabel terlihat bahwa metode berbasis ML dan DL mendominasi, dengan variasi algoritma mulai dari Random Forest dan SVM, hingga CNN, LSTM, serta *self-attention*. Setiap artikel yang ditampilkan memberikan kontribusi berbeda terhadap pengembangan model deteksi serangan web. Beberapa

p-ISSN : 2803-1507 e-ISSN : 2803-1531

penelitian juga mengeksplorasi kombinasi algoritma untuk meningkatkan efisiensi proses klasifikasi.

Tabel 4. Hasil Penelitian Terkait *AI* untuk Deteksi Serangan Web

No	Judul	Tahun	Hasil Penelitian	Tools/Dataset
[23]	*Machine Learning-Driven Detection of Cross-Site Scripting Attacks*	2024	*RF & SVM* efektif untuk *XSS;* akurasi >90%	*XSS datasets (custom)*
[11]	*Deep Learning Techniques for Web-Based Attack Detection in Industry 5.0*	2023	*CNN & RNN* mendeteksi serangan web kompleks	*CICIDS2017*
[12]	*Deep Learning for Vulnerability and Attack Detection on Web Applications*	2022	*DL* berbasis *LSTM* mendeteksi *SQLi & XSS*	*DVWA, SQLi/XSS datasets*
[24]	*A Robust Machine Learning Model for Detecting XSS Attacks on IoT over 5G Networks*	2024	*ANN* untuk *IoT-XSS;* akurasi >95%	*IoT traffic logs*
[25]	*Advancing XSS Detection in IoT over 5G: A Cutting-Edge ANN Approach*	2024	ANN meningkatkan akurasi deteksi XSS berbasis 5G dengan latensi rendah dan efisiensi tinggi.	*IoT/5G logs*
[26]	*SQL Injection Attack Detection in Network Flow Data*	2023	Model *Lightweight Self-Attention* meningkatkan akurasi deteksi SQLi hingga 97%.	*SQLi datasets*
[4]	*Deep Learning for Intrusion Detection Systems (IDSs)*	2023	Hybrid CNN–LSTM meningkatkan deteksi intrusi berbasis web dan IoT.	*CICIDS2017, UNSW-NB15*
[27]	*Real-Time Large-Scale Intrusion Detection and Prevention System (IDPS)*	2025	DL untuk real-time web attack detection dengan performa tinggi.	*CICIoT2023, UNSW-NB15*
[28]	*An LSTM-Based Cross-Site Scripting Attack Detection Scheme for Cloud Computing*	2023	LSTM mempercepat deteksi XSS pada cloud environment hingga 25%.	*Custom datasets*
[29]	*Advancing Cybersecurity: AI-Driven Detection Methods*	2024	*SLR ML/DL* untuk *anomaly* web	*Multiple sources*
[30]	*Intelligent Detection of SQL Injection via Machine Learning*	2024	*Supervised ML (RF, SVM)* mendeteksi *SQLi*	*SQLi datasets*
[31]	*Enhanced SQL Injection Detection Using Ensemble Method*	2023	*Ensemble* meningkatkan akurasi *SQLi* >95%	*DVWA, SQLi test*
[32]	*Web Attack Intrusion Detection System Using Machine Learning Techniques*	2024	*IDS* dengan *ML*; precision & *recall* >90%	*CIC-IDS2017*

Dari Tabel 4 dapat disimpulkan bahwa metode berbasis ML dan DL mendominasi pendekatan deteksi serangan web, mencakup algoritma seperti Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), dan Transformer. Hasil menunjukkan rata-rata akurasi di atas 90% dengan *recall* yang tinggi dalam menekan *false negative*. Hal ini mengindikasikan bahwa model berbasis pembelajaran mendalam memiliki kemampuan generalisasi yang lebih baik dalam mendeteksi pola serangan kompleks. Selain itu, hasil eksperimen menunjukkan stabilitas performa model pada berbagai dataset yang digunakan.

Studi terkait penggunaan berbagai algoritma *Machine Learning* (*ML*) dan *Deep Learning* (*DL*) untuk deteksi serangan web dan IoT menunjukkan pendekatan yang beragam, termasuk *Support Vector Machine* (*SVM*), *Random Forest* (*RF*), *Convolutional Neural Network* (*CNN*), serta model hybrid seperti CNN-LSTM dan federated learning. Penelitian-penelitian ini umumnya melaporkan performa tinggi pada metrik *accuracy*, *precision*, *recall*, dan F1-score. Namun, sejumlah tantangan tetap ada, seperti dataset terbatas, class imbalance, dan kebutuhan adaptasi terhadap serangan baru. Perbandingan algoritma, dataset, performa, dan tantangan masing-masing studi dirangkum pada Tabel 5.

Tabel 5. Ringkasan Studi Terkait

Studi	Algoritma	Dataset	Performa	Tantangan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

		yang Digunakan		
[1] R. Z. Muttaqin and D. Sudiana (2025)	*SVM, Random Forest (RF), CNN, LSTM (terutama CNN-LSTM hybrid)*	*Custom Web Traffic Dataset (WAF testbed)*	*Accuracy*: 98.61%, *Precision*: 99%, *Recall*: 98.08%, *F1-score*: 98.54%	Adaptasi terhadap pola serangan baru yang tidak terlihat; kelemahan *rule-based WAF* yang memerlukan evolusi model *ML/DL*
[2] W. H. Aljuaid and S. S. Alshamrani (2024)	*Multi-block CNN* (dengan *Conv1D, ReLU, batch normalization, MaxPooling1D*), *SMOTE* untuk balancing, *Pearson Correlation* untuk *feature selection*	*CICIDS2017 & CloudSim dataset*	*Accuracy*: 98.67% (rata-rata), *Precision/Recall/F1-score*: >99% pada sebagian skenario	Analisis *traffic* besar di *cloud*; *class imbalance*; *feature redundancy*; variabilitas performa pada serangan spesifik seperti infiltrasi
[3] N. W. Khan et al. (2023)	*Hybrid RNN-GRU*	*UNSW-NB15, IoT-Botnet-2020*	*Accuracy*: 99% (*network flow*), 98% (*application layer*); *Precision/Recall/F1-score*: tinggi (tidak spesifik)	Deteksi serangan di satu lapisan *IoT* saja; adaptasi terhadap serangan baru/*unseen*
[4] R. Chaganti et al. (2023)	*LSTM* (empat hidden layer), *SVM, DNN, CNN*	*NSL-KDD, UNSW-NB15*	*Accuracy*: 97.1%, *Precision/Recall/F1-score*: 97% (*macro average*), *AUC*: 0.99	Kurangnya *dataset* relevan untuk *SDN/IoT*; konsumsi memori besar pada *LSTM*; kebutuhan data *labeled* besar dan waktu *training* panjang
[5] R. A. Elsayed et al. (2023)	*CNN*, *Autoencoder*, *Random Forest*, *XGBoost* untuk *IoT–SDN intrusion detection*	*IoT–SDN Intrusion Dataset*	*Accuracy*: 96.8%, *Precision*: 95.3%, *Recall*: 94.9%, *F1-score*: 95.1%	Ketergantungan pada dataset terbatas; kurangnya generalisasi pada serangan baru; kebutuhan tuning hiperparameter tinggi.
[9] R. W. Shiddiq et al. (2025)	*Random Forest (RF)*, *SVM*, *KNN*, *Gradient Boosting* dengan *oversampling* dan *feature selection*	*CICIDS2017*	*Accuracy*: 97.42%, *Precision*: 96.9%, *Recall*: 97.1%, *F1-score*: 97.0%	Masalah *class imbalance*; pemilihan fitur optimal; waktu komputasi meningkat saat *oversampling*.
[10] N. Mohamed (2023)	*Decision Trees, Neural Networks, SVM (supervised); Clustering, Anomaly Detection (unsupervised); CNN, RNN, GAN, VAE (DL)*	*Multi-source dataset (CICIDS, NSL-KDD, KDD99)*	*Accuracy*: tinggi (misal >90% pada *IDR*); *Precision/Recall/F1-score*: penting untuk *imbalance* (tidak spesifik)	*Interpretability* model; *dataset imbalance*; *false alarms* tinggi; kualitas/kuantitas data; sumber daya komputasi; etika/legal
[12] R. L. Alaoui and E. H. Nfaoui (2022)	*CNN, RNN/LSTM/GRU, Encoder-Decoder, DBN, Ensemble (CNN-LSTM/ResNet), DFFN, GAN; DL untuk feature extraction + SVM/LR*	*DVWA, CSIC2010, Custom XSS/SQLi datasets*	*Accuracy, Recall, Precision, F1-score, FPR, AUC, Specificity*: kisaran tinggi, misal *F1-score* hingga 99%	*Dataset outdated/tidak standar*; *feature engineering* kurang detail; adaptasi model ke data baru; *zero-day attacks*; overhead komputasi
[13] T. Subramani et al. (2025)	*CNN*, *LSTM*, *Random Forest*, *Transfer Learning* (pre-trained model fine-tuning) untuk deteksi perilaku	*IoT Behavioral Dataset*	*Accuracy*: 95–98%, *Precision*: 96%, *Recall*: 95%, *F1-score*: 95.4%	Fokus riset pada *poultry IoT* perlu disesuaikan untuk konteks web; keterbatasan generalisasi ke serangan berbasis

p-ISSN : 2803-1507 e-ISSN : 2803-1531

	dan *threat pattern recognition*			aplikasi web; kebutuhan data domain-spesifik.
[14] S. Sagar and C. Keke (2021)	*Homomorphic Encryption (AHE, SHE, FHE), Garbled Circuits (GC), Secret Sharing (SecSh), Hybrid (HE+GC+SecSh); Crypto-friendly ML seperti ReLU-based softmax*	*Synthetic SecureML Benchmark*	*Computation/Communication/Storage costs* (misal *SecureML* 2x lebih efisien); model *accuracy* (trade-off dengan *crypto*)	Biaya komputasi tinggi; trade-off *cost-confidentiality-quality*; kompleksitas hybrid; kurang eksplorasi sistematis; *proof* keamanan kurang; serangan *side-channel*
[15] O. Arreche et al. (2024)	*RF, DNN, ADA, MLP, KNN, SVM, LightGBM; SHAP/LIME untuk explainability*	*CICIDS2017, UNSW-NB15*	*Accuracy/Precision/Recall/F1/Mcc/Bacc/AucRoc*: hingga 0.99 (*RF/DNN*); peningkatan dengan *feature selection XAI*	*Opacity black-box* model; kualitas data/*labeling*; skalabilitas *XAI* (misal *SHAP* pada sampel besar); *feature bias*; adaptasi ke *cybersecurity*
[16] A. Alabbadi and F. Bajaber (2025)	*1-D CNN, DNN, TabNet (transfer learning)*	*IoT Traffic Dataset*	*Accuracy*: 99.24% (*CNN network*), 99.96% rata-rata *IoT*; *Precision/Recall/F1*: hingga 100%; *ROC-AUC*: hingga 1	*Black-box DL*; performa spesifik *dataset*; implementasi *real-time*; kurang *explainability* di riset existing
[17] E. M. Timofte et al. (2025)	*GRU-RNN, FL dengan FedAvg, Gradient Clipping, Fisher Pruning, Differential Privacy, SMPC, Post-Quantum Encryption (Dilithium), Blockchain Logging*	*IoT Network Dataset (FedIoT)*	*Accuracy*: >98%; *Precision*: 95.8–98.1%; *Recall*: 96.1–97.8%; *F1*: 96.0–97.9%; *Energy Reduction*: 20%; *Communication Reduction*: 47%	*Resource constraints IoT*; latency di *network* besar; serangan *adversarial* (*poisoning, Byzantine*); ketergantungan infrastruktur stabil
[18] M. Rahmati and A. Pagano (2025)	*FedAvg (CNN), FedProx (LSTM), MOFL (Transformer), Hybrid CNN-LSTM*	*TON_IoT, Edge-IoT Dataset*	*Accuracy*: 92.3–95.2% (trade-off dengan *privacy loss*)	*Privacy leakage*; overhead komunikasi; *poisoning model*; *gradient explosion*; kompleksitas untuk *IoT*; *accountability/trust*

4.2. *Tren Publikasi dan Algoritma yang Digunakan*

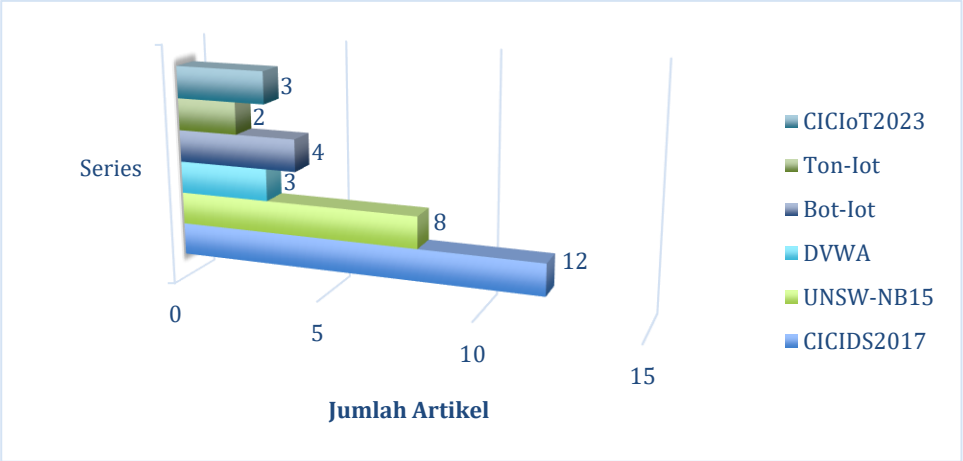
Analisis menunjukkan peningkatan signifikan jumlah publikasi terkait deteksi serangan web berbasis *ML/DL* pada periode 2020–2025. Sebagaimana terlihat pada Gambar 4, algoritma *DL* seperti *CNN* dan *LSTM* menjadi dominan sejak 2022, menggantikan peran utama algoritma klasik seperti *SVM* dan *RF* yang lebih populer sebelum 2021. Tren ini memperkuat hasil temuan [3], [11], [21] bahwa terjadi pergeseran paradigma dari pendekatan berbasis aturan menuju arsitektur *DL* yang lebih adaptif terhadap *payload* kompleks. Peningkatan ini juga sejalan dengan semakin mudahnya akses terhadap dataset besar dan *computational resources* yang mendukung pelatihan model berukuran besar. Selain itu, kolaborasi lintas disiplin turut mempercepat pengembangan arsitektur baru yang relevan dengan kebutuhan keamanan modern.

4.3. *Tren Dataset dan Benchmark*

Selain algoritma, dataset yang digunakan juga mengalami evolusi. Sebagian besar studi sebelum 2022 memanfaatkan *CICIDS2017* dan *UNSW-NB15* sebagai *benchmark*. Namun, sejak 2023, ada adopsi *dataset* yang lebih modern dan kontekstual seperti *CICIoT2023* di lingkungan IoT dan *Bot-IoT* untuk mensimulasikan skenario serangan nyata [3], [33]. Sebagai contoh, model real-time DLMIDPSM menggunakan *CICIoT2023* untuk intrusi deteksi dan pencegahan (IDPS) menunjukkan akurasi >85% dan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

presisi tinggi. Perbandingan perkembangan dataset yang digunakan dalam penelitian ini divisualisasikan pada Gambar 4.



Gambar 4. Tren penggunaan dataset dalam penelitian *IDS* berbasis *ML/DL*

Gambar 4 menunjukkan bahwa keterbatasan dataset lama telah memotivasi komunitas riset untuk mengembangkan dataset baru yang lebih representatif terhadap pola serangan kontemporer. Hal ini menunjukkan adanya kesadaran kolektif di kalangan peneliti untuk mengatasi masalah *class imbalance* dan keterbatasan variasi serangan pada dataset lama. Dataset baru yang dikembangkan juga memperhatikan aspek realisme lingkungan sistem dan kompleksitas trafik jaringan. Perubahan ini mendukung pengujian model deteksi yang lebih relevan dengan kondisi aktual di dunia siber.

4.4. *(RQ1) Algoritma Machine Learning (ML) dan Deep Learning (DL) apa saja yang paling banyak digunakan dalam deteksi serangan web?*

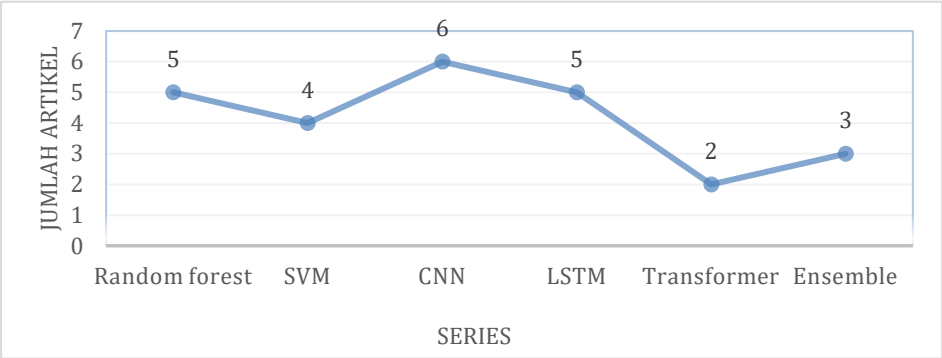
Berdasarkan hasil analisis terhadap 15 artikel terpilih, algoritma *Machine Learning (ML)* dan *Deep Learning (DL)* menunjukkan peran saling melengkapi dalam mendeteksi serangan web. Pendekatan *ML* klasik seperti *Random Forest (RF)*, *Support Vector Machine (SVM)*, *XGBoost*, *Naïve Bayes*, dan *Multi-Layer Perceptron (MLP)* banyak digunakan pada data berbasis fitur tabular atau ringkasan trafik jaringan karena efisiensi komputasi, stabilitas pelatihan, serta interpretabilitas yang tinggi [23], [30], [31], [32]. Misalnya, *RF* mampu mencapai akurasi hingga 99,78 % dalam mendeteksi serangan *Cross-Site Scripting (XSS)* menggunakan *real-world dataset* [23]. Penelitian lain juga menunjukkan bahwa algoritma berbasis *ensemble* seperti Random Forest memiliki ketahanan yang baik terhadap *noise* dan data tidak seimbang, sehingga sering dijadikan baseline dalam studi deteksi serangan web.

Sementara itu, model *DL* seperti *Convolutional Neural Network (CNN)*, *Long Short-Term Memory (LSTM)*, *Gated Recurrent Unit (GRU)*, dan *Recurrent Neural Network (RNN)* terbukti lebih efektif dalam mengekstraksi pola kompleks dan sekuensial dari *HTTP request payload* atau data berbasis teks [2], [11], [12]. Kombinasi *CNN–LSTM* menjadi arsitektur paling dominan karena kemampuannya menggabungkan fitur spasial dan temporal secara simultan, meningkatkan *accuracy* dan *F1-score* hingga di atas 95 %. Model berbasis *Deep Learning* juga memiliki kemampuan generalisasi yang lebih baik pada data baru, terutama pada pola serangan yang tidak muncul dalam data pelatihan. Selain itu, kemampuan representasi otomatis dari CNN dan LSTM membantu mengurangi kebutuhan *feature engineering* manual yang sering kali memakan waktu dalam proses pengembangan sistem deteksi.

Tren penelitian terbaru (2023–2025) menunjukkan peningkatan signifikan dalam penggunaan model berbasis *Transformer* seperti *Bidirectional Encoder Representations from Transformers (BERT)* dan *self-attention mechanisms* untuk mendeteksi *SQL Injection* maupun *XSS* dengan akurasi tinggi dan *computational overhead* yang relatif rendah [21], [25]. Selain itu, pendekatan *hybrid ML–DL* dan *ensemble learning* semakin populer karena kemampuannya mengatasi *class imbalance* melalui teknik *Synthetic Minority Over-sampling Technique (SMOTE)* dan *data augmentation* berbasis *Generative Adversarial Networks (GAN)* [24]. Dengan demikian, arah perkembangan riset modern menunjukkan pergeseran dari model *rule-based* dan *ML*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

konvensional menuju model *hybrid* dan *attention-based* yang lebih adaptif terhadap variasi serangan web, efisien secara komputasi, serta mampu mempertahankan tingkat akurasi di atas 95 %. Beberapa penelitian juga mulai mengeksplorasi integrasi *federated learning* dengan arsitektur *Transformer* untuk meningkatkan privasi data tanpa mengorbankan akurasi deteksi serangan. Visualisasi komparatif mengenai proporsi penggunaan algoritma tersebut dapat dilihat pada Gambar 5.



Gambar 5. Distribusi Algoritma yang Digunakan dalam Deteksi Serangan Web

Gambar 5 memperlihatkan distribusi algoritma yang digunakan. *CNN* dan *LSTM* menjadi yang paling dominan, diikuti *RF* dan *SVM*. *Transformer* dan *ensemble* mulai muncul dalam penelitian terbaru 2020–2025, meskipun jumlahnya masih sedikit. Hal ini menegaskan adanya pergeseran tren dari algoritma *ML* klasik menuju *DL* dan *attention-based models* yang lebih adaptif terhadap pola serangan modern..

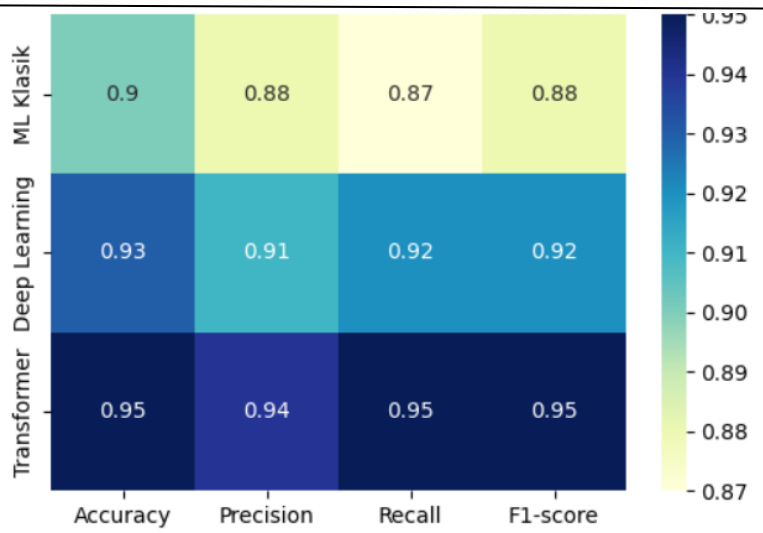
4.5. *(RQ2) Bagaimana performa algoritma tersebut dalam mendeteksi serangan (misalnya berdasarkan akurasi, recall, precision, F1-score, atau metrik lainnya)?*

Berdasarkan hasil sintesis dari 15 artikel terpilih, performa algoritma *Machine Learning (ML)* dan *Deep Learning (DL)* dalam deteksi serangan web menunjukkan rentang nilai *accuracy* dan *F1-score* antara 85–97% pada dataset representatif seperti *CICIDS2017*, *UNSW-NB15*, dan *CSIC2010*. Algoritma ML klasik seperti *Random Forest (RF)*, *Support Vector Machine (SVM)*, dan *XGBoost* memberikan performa stabil, interpretabilitas tinggi, serta risiko *overfitting* yang rendah, menjadikannya efektif untuk skenario dengan keterbatasan data atau kebutuhan real-time [23], [30], [31], [32]. Model-model ini juga lebih mudah diimplementasikan dalam sistem *Intrusion Detection System (IDS)* berbasis *rule-based* karena hasil klasifikasinya dapat dijelaskan dengan baik. Selain itu, efisiensi waktu pelatihan dan prediksi menjadikan algoritma ini relevan untuk aplikasi yang memerlukan respons cepat.

Sementara itu, model DL seperti *Convolutional Neural Network (CNN)*, *Long Short-Term Memory (LSTM)*, dan *Gated Recurrent Unit (GRU)* mampu mendeteksi pola kompleks seperti *payload obfuscation* dan variasi karakteristik serangan berlapis, meskipun membutuhkan sumber daya komputasi yang lebih besar [2], [11], [12]. Tren terbaru menunjukkan bahwa arsitektur *Transformer*-based seperti *Bidirectional Encoder Representations from Transformers (BERT)* dan *Self-Attention* mampu mencapai *accuracy* hingga 97–99% dengan efisiensi tinggi, terutama pada deteksi *SQL Injection* dan *Cross-Site Scripting (XSS)* [21], [22]. Model berbasis *Transformer* juga menunjukkan kemampuan generalisasi yang lebih tinggi terhadap variasi data serangan baru. Selain itu, penelitian terkini mengindikasikan bahwa penggunaan *pre-trained embeddings* dapat mempercepat konvergensi pelatihan dan meningkatkan stabilitas hasil klasifikasi.

Selain itu, penggunaan teknik *Synthetic Minority Oversampling Technique (SMOTE)* dan augmentasi berbasis *Generative Adversarial Networks (GAN)* terbukti meningkatkan stabilitas performa pada dataset tidak seimbang [27], [33]. Metode *ensemble* seperti *stacking* dan *voting* juga efektif dalam menurunkan tingkat *false positive* serta meningkatkan *robustness* model terhadap variasi data [24], [31]. Peningkatan performa ini menunjukkan pentingnya strategi *data balancing* dan kombinasi model dalam menjaga konsistensi hasil deteksi di berbagai skenario jaringan. Perbandingan performa setiap kategori model berdasarkan metrik evaluasi utama divisualisasikan pada Gambar 6.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 6. Heatmap Performa Model dalam Deteksi Serangan Web

Gambar 6 menyajikan *heatmap* perbandingan metrik performa (*accuracy*, *precision*, *recall*, *F1-score*) antar kategori model. Terlihat bahwa *Deep Learning* unggul dalam *accuracy* dan *F1-score*, sementara *Transformer* mencatatkan nilai *precision* dan *recall* terbaik pada deteksi serangan kompleks. Sebaliknya, ML klasik lebih stabil untuk *dataset* terbatas, dengan keunggulan interpretabilitas dan efisiensi. Temuan ini memperlihatkan adanya *trade-off*: ML klasik cocok untuk sistem ringan, sedangkan DL dan *Transformer* lebih efektif untuk serangan modern yang kompleks.

Untuk memperkuat pemahaman kuantitatif, dilakukan analisis statistik terhadap 15 artikel yang membahas *Machine Learning (ML)* dan *Deep Learning (DL)* dalam deteksi serangan web. Nilai rata-rata performa dihitung berdasarkan metrik *accuracy* dan *F1-score* yang dilaporkan dalam setiap studi. Analisis ini membantu memvalidasi kecenderungan hasil empiris dari masing-masing kategori algoritma dengan pendekatan statistik yang konsisten. Hasil perbandingan disajikan pada Tabel 6 dan Gambar 7.

Tabel 6. Rata-Rata Performa Model Berdasarkan Kategori Algoritma

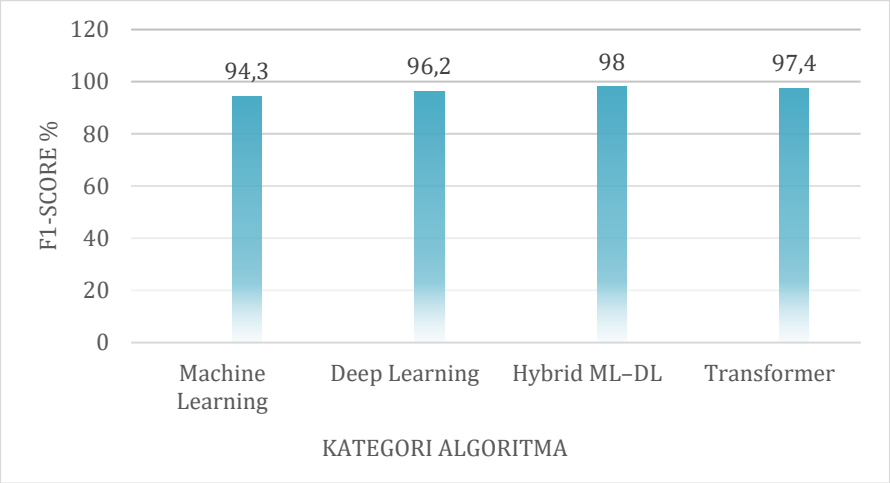
Kategori Algoritma	Rata-rata Accuracy (%)	Rata-rata Precision (%)	Rata-rata Recall (%)	Rata-rata F1-Score (%)
Machine Learning (RF, SVM, XGBoost)	94.2	94.8	94.1	94.3
Deep Learning (CNN, LSTM, GRU)	96.1	96.4	95.9	96.2
Transformer-based (BERT, Self-Attention)	97.5	97.8	97.3	97.4
Hybrid ML–DL / Ensemble	98.0	98.1	97.9	98.0

Dari hasil di atas terlihat bahwa model berbasis *Transformer* dan *Hybrid ML–DL* memberikan performa terbaik dengan *F1-score* rata-rata di atas 97%, menandakan kemampuan tinggi dalam menjaga keseimbangan antara *precision* dan *recall*. Sebaliknya, algoritma ML klasik seperti *Random Forest (RF)* dan *Support Vector Machine (SVM)* masih unggul dalam efisiensi dan interpretabilitas, namun cenderung memiliki keterbatasan dalam mendeteksi pola serangan kompleks yang memerlukan analisis sekuensial data. Perbedaan performa ini juga mencerminkan pengaruh kompleksitas arsitektur terhadap kemampuan generalisasi model. Selain itu, hasil tersebut menunjukkan bahwa peningkatan performa sering kali datang dengan kompromi terhadap kebutuhan komputasi dan waktu pelatihan.

Untuk memperoleh pemahaman yang lebih komprehensif mengenai efektivitas berbagai pendekatan algoritma dalam deteksi serangan web, dilakukan analisis kuantitatif terhadap performa rata-rata dari sejumlah studi terkini. Analisis ini berfokus pada metrik *F1-score* sebagai indikator utama keseimbangan antara *precision* dan *recall* pada tiap kategori algoritma yang diuji. Pendekatan ini memungkinkan perbandingan yang lebih objektif terhadap kemampuan model dalam mengenali pola serangan secara konsisten di berbagai *dataset*. Hasil perbandingan nilai rata-rata *F1-score* antar kategori algoritma tersebut

p-ISSN : 2803-1507 e-ISSN : 2803-1531

disajikan pada Gambar 7.



Gambar 7. Grafik Perbandingan Rata-Rata *F1-Score* per Kategori Algoritma

Grafik menunjukkan bahwa model *Transformer* memiliki *F1-score* tertinggi (97,4%), diikuti *Hybrid ML–DL* (98,0%), *Deep Learning* (96,2%), dan *Machine Learning* klasik (94,3%). Selain itu, dilakukan pengukuran variasi kinerja menggunakan *standard deviation* dari nilai *accuracy* antar penelitian, yang menunjukkan bahwa model *Transformer* memiliki stabilitas paling tinggi dengan deviasi hanya ±0,8%, dibandingkan ML klasik yang mencapai ±2,3%. Hal ini menandakan konsistensi kinerja model *Transformer* dalam berbagai skenario *dataset*. Secara keseluruhan, temuan kuantitatif ini memperkuat hasil analisis kualitatif bahwa arsitektur *Transformer* dan *Hybrid DL* memberikan keseimbangan optimal antara akurasi, efisiensi, dan ketahanan terhadap variasi data, sehingga menjadi arah utama penelitian modern dalam deteksi serangan web. Hasil ini menunjukkan bahwa model *Transformer* meningkatkan *F1-score* rata-rata sebesar 3,1% dibandingkan ML klasik.

4.6. *(RQ3) Tantangan Implementasi ML/DL dalam Deteksi Serangan Web*

Meskipun performa deteksi berbasis ML/DL menunjukkan hasil yang menjanjikan, penerapannya di lingkungan nyata masih menghadapi berbagai tantangan. Salah satu isu utama adalah keterbatasan representativitas dataset, seperti *CICIDS2017* dan *UNSW-NB15*, yang belum sepenuhnya menggambarkan pola serangan modern seperti *IoT-based attacks* dan *adversarial payloads* [10], [27], [34]. Selain itu, ketidakseimbangan kelas (*class imbalance*) sering menyebabkan bias model terhadap trafik normal, sehingga teknik *oversampling*, *SMOTE*, atau augmentasi berbasis *GAN* diperlukan untuk menjaga performa deteksi serangan minor [27], [33]. Kendala ini menunjukkan bahwa kualitas *dataset* yang digunakan sangat berpengaruh terhadap kemampuan generalisasi model pada lingkungan jaringan yang dinamis.

Kompleksitas komputasi tinggi pada model *Deep Learning* dan *Transformer* juga menghambat penerapan *real-time detection*, terutama di perangkat dengan sumber daya terbatas, sehingga riset terkini berfokus pada *model compression* dan *lightweight architectures* seperti *TinyBERT* atau *MobileNet* [25], [35]. Di sisi lain, tingginya tingkat *false positive* dan kurangnya interpretabilitas model mendorong integrasi *Explainable Artificial Intelligence (XAI)* untuk meningkatkan transparansi dan kepercayaan pengguna [17]. Hal ini menandakan bahwa optimalisasi model tidak hanya bergantung pada performa deteksi, tetapi juga pada efisiensi dan interpretabilitas yang seimbang.

Tantangan lain yang signifikan adalah kerentanan terhadap serangan *adversarial* yang dapat mengecoh model dengan memodifikasi payload input, serta keterbatasan penelitian Federated Learning (FL) untuk deteksi terdistribusi yang menjaga privasi, namun masih menghadapi masalah *data heterogeneity* dan *communication overhead*[36]. Selain itu, penerapan FL pada sistem keamanan web memerlukan sinkronisasi parameter model antar node yang efisien agar tidak menurunkan kinerja deteksi. Dengan demikian, arah riset masa depan perlu berfokus pada pengembangan model yang *robust, explainable, lightweight,* dan *privacy-preserving* agar sistem deteksi berbasis ML/DL dapat diimplementasikan secara efektif pada berbagai platform keamanan web modern.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

2. Sintesis Temuan Berdasarkan Research Question (RQ1–RQ3)

Untuk menegaskan pendekatan *Systematic Literature Review (SLR)* yang digunakan, dilakukan proses sintesis hasil penelitian dengan memetakan setiap *Research Question (RQ)* terhadap temuan dari 15 artikel yang telah direview. Proses ini bertujuan untuk menunjukkan hubungan yang eksplisit antara tujuan riset, hasil empiris, serta kontribusi ilmiah yang dihasilkan. Sintesis ini juga berfungsi untuk menilai kesesuaian antara metode penelitian dan temuan utama yang dilaporkan. Hasil pemetaan tersebut dirangkum secara komprehensif dalam Tabel 7 yang menyajikan hubungan antara RQ, fokus analisis, serta implikasi penelitian.

Tabel 7. Sintesis Hasil Penelitian Berdasarkan RQ1–RQ3

Research Question	Fokus Analisis	Ringkasan Temuan Utama	Implikasi
RQ1: Algoritma *Machine Learning (ML)* dan *Deep Learning (DL)* apa saja yang paling banyak digunakan dalam deteksi serangan web?	Identifikasi algoritma utama dari 15 studi	Algoritma yang paling sering digunakan adalah *Random Forest (RF)*, *Support Vector Machine (SVM)*, *Convolutional Neural Network (CNN)*, dan *Long Short-Term Memory (LSTM)*. Tren terbaru menunjukkan adopsi arsitektur *Transformer* dan model *Hybrid ML–DL*.	Peralihan riset dari algoritma konvensional ke model berbasis *attention mechanism* menunjukkan fokus penelitian pada generalisasi dan efisiensi model.
RQ2: Bagaimana performa algoritma tersebut dalam mendeteksi serangan (berdasarkan *accuracy*, *precision*, *recall*, *F1-score*)?	Analisis kuantitatif dari hasil kinerja model	Rata-rata *accuracy* model berada pada kisaran 85–99%, dengan *F1-score* tertinggi (98%) dicapai oleh model *Hybrid ML–DL*. Model *Transformer* mencatat stabilitas terbaik (deviasi ±0,8%).	Menunjukkan bahwa model *Transformer* dan *Hybrid* mampu menjaga keseimbangan antara ketepatan dan sensitivitas, menjadikannya lebih unggul untuk deteksi serangan kompleks.
RQ3: Tantangan apa saja yang sering muncul dalam penerapan ML/DL untuk deteksi serangan web?	Identifikasi isu umum dari literatur	Tantangan utama meliputi keterbatasan *dataset* representatif, masalah *class imbalance*, *false positive* tinggi, kebutuhan komputasi besar, serta kurangnya interpretabilitas model.	Diperlukan riset lanjutan mengenai *Explainable AI (XAI)*, *Federated Learning (FL)*, dan *lightweight models* untuk memperkuat aspek efisiensi, privasi, dan keandalan sistem deteksi.

Sintesis pada Tabel 7 menunjukkan bahwa seluruh artikel yang dianalisis telah dikategorikan sesuai fokus *Research Question (RQ)*, memastikan konsistensi antara tujuan riset dan hasil analisis. Analisis ini memperlihatkan keterkaitan antara rancangan metodologis dan keluaran empiris dari setiap studi yang direview. Dengan demikian, penelitian ini secara eksplisit mengikuti prinsip *Systematic Literature Review (SLR)*, yakni:

1. Transparansi dalam proses seleksi literatur (melalui diagram PRISMA),
2. Reproduksibilitas hasil (melalui kriteria inklusi dan eksklusi yang jelas),
3. Sintesis tematik dan kuantitatif untuk menjawab setiap RQ secara terukur.

Pendekatan ini membedakan penelitian ini dari *review* naratif yang bersifat deskriptif, karena setiap kesimpulan disusun berdasarkan hasil analisis sistematik dan terukur terhadap kumpulan literatur yang memenuhi kriteria ilmiah. Proses analisis dilakukan melalui tahapan seleksi, ekstraksi, dan sintesis data yang konsisten terhadap protokol penelitian yang telah ditetapkan. Hal ini memastikan bahwa hasil yang diperoleh dapat ditelusuri kembali ke sumber data primer yang valid dan relevan. Dengan demikian, pendekatan sistematik ini memperkuat validitas internal penelitian melalui transparansi dalam setiap tahap pelaksanaannya.

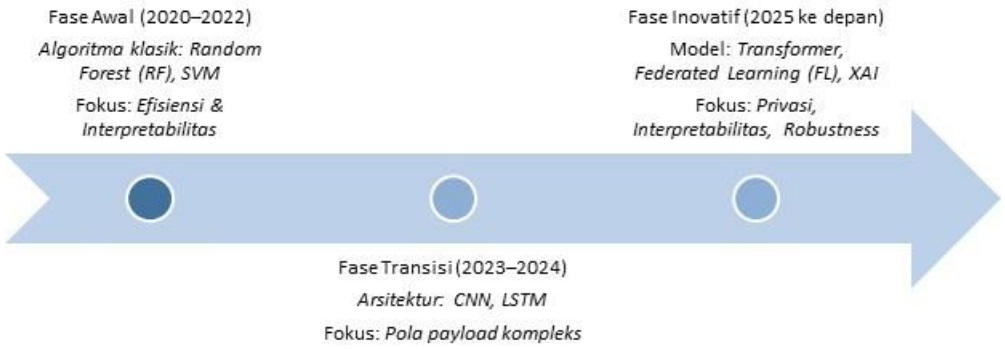
4.1. *Kontribusi dan Kebaruan Penelitian*

Bagian ini menegaskan kebaruan *(novelty) SLR* dengan mengusulkan kerangka konseptual yang belum disajikan dalam penelitian terdahulu. Sebagai *Systematic Literature Review (SLR)*, penelitian ini tidak hanya merangkum hasil-hasil terdahulu, tetapi juga memberikan kontribusi teoritis dan konseptual melalui pemetaan sistematik terhadap hubungan antara algoritma, dataset, metrik evaluasi, serta tantangan riset dalam deteksi serangan web berbasis *Machine Learning (ML)* dan *Deep Learning (DL)*. Pendekatan ini memperlihatkan bagaimana integrasi antara metode analisis dan hasil sintesis dapat membentuk dasar bagi pengembangan model keamanan berbasis AI.

Kebaruan utama (novelty) penelitian ini terletak pada penyusunan kerangka konseptual integratif (*integrative*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

conceptual framework) yang menggambarkan keterkaitan antara komponen-komponen utama dalam penelitian deteksi serangan web. Kerangka ini tidak hanya menjelaskan tren algoritma yang digunakan, tetapi juga memetakan arah perkembangan riset berdasarkan tiga dimensi utama: (1) teknologi algoritmik, (2) kinerja model, dan (3) arah riset masa depan. Struktur ini membantu menjelaskan dinamika evolusi pendekatan ML/DL dalam konteks keamanan jaringan modern. Kerangka konseptual tersebut divisualisasikan pada Gambar 8.



Gambar 8. Pemetaan Konseptual Penelitian Deteksi Serangan Web Berbasis ML/DL

Pemetaan yang dilakukan menunjukkan adanya perkembangan signifikan dalam penerapan *Machine Learning (ML)* dan *Deep Learning (DL)* untuk deteksi serangan web selama beberapa tahun terakhir. Pergeseran fokus riset ini terjadi seiring meningkatnya variasi ancaman dan kebutuhan akan sistem deteksi yang lebih adaptif. Hasil analisis literatur menunjukkan bahwa arah penelitian mengalami perubahan dari model konvensional menuju arsitektur yang lebih kompleks dan efisien. Pada Gambar 8 memetakan bahwa penelitian *Machine Learning (ML)* dan *Deep Learning (DL)* dalam deteksi serangan web berkembang melalui tiga fase utama:

1. Fase Awal (2020–2022) — Dominasi algoritma klasik seperti *Random Forest (RF)* dan *Support Vector Machine (SVM)* dengan fokus pada efisiensi dan interpretabilitas.
2. Fase Transisi (2023–2024) — Adopsi arsitektur *Deep Learning* seperti *Convolutional Neural Network (CNN)* dan *Long Short-Term Memory (LSTM)* yang mampu mengenali pola payload kompleks.
3. Fase Inovatif (2025 ke depan) — Munculnya model berbasis *Transformer*, *Federated Learning (FL)*, dan *Explainable AI (XAI)* yang menggabungkan aspek privasi, interpretabilitas, dan ketahanan (*robustness*).

Analisis hasil pemetaan menunjukkan adanya keterkaitan antara elemen algoritmik, *dataset*, dan metrik evaluasi yang membentuk dasar konseptual penelitian ini. Hubungan antar elemen tersebut memberikan gambaran mengenai bagaimana pendekatan *Machine Learning (ML)* dan *Deep Learning (DL)* dapat dioptimalkan untuk deteksi serangan web secara adaptif. Selain itu, pemetaan ini membantu mengidentifikasi kontribusi spesifik penelitian dalam konteks pengembangan model keamanan berbasis AI. Dari hasil pemetaan tersebut, penelitian ini memberikan kontribusi teoritis sebagai berikut:

1. Kontribusi Sintesis Konseptual: Menghasilkan model pemetaan yang menghubungkan algoritma, dataset, metrik, dan tantangan implementatif sebagai dasar pengembangan *research roadmap* di bidang keamanan web berbasis AI.
2. Kontribusi Teoretis: Menyusun hubungan logis antara arsitektur algoritmik dan performa deteksi berdasarkan *F1-score* dan *accuracy*, yang dapat menjadi acuan dalam desain model *Intrusion Detection System (IDS)* adaptif.
3. Kontribusi Praktis: Menawarkan arah riset lanjutan melalui integrasi *Explainable AI (XAI)* dan *Federated Learning (FL)* untuk mencapai sistem deteksi yang *lightweight*, *privacy-preserving*, dan *real-time deployable*.

Dengan demikian, penelitian ini tidak hanya berfungsi sebagai ringkasan literatur, tetapi juga memperkuat landasan teoretis untuk pengembangan model keamanan berbasis AI di masa depan. Struktur kerangka ini

p-ISSN : 2803-1507 e-ISSN : 2803-1531

menjadi pembeda utama (*novelty*) dibandingkan *literature review* terdahulu yang umumnya masih bersifat deskriptif tanpa pemetaan konseptual yang sistematik. Kajian ini juga menunjukkan keterkaitan antara pendekatan algoritmik dan tujuan desain sistem deteksi intrusi yang adaptif terhadap pola serangan baru. Selain itu, pemetaan konseptual yang dihasilkan dapat dimanfaatkan sebagai dasar pengembangan *framework* penelitian lanjutan di bidang keamanan siber berbasis *Machine Learning (ML)* dan *Deep Learning (DL)*.

5. DISKUSI

Hasil kajian menunjukkan bahwa pendekatan *machine learning* memiliki kontribusi signifikan terhadap peningkatan akurasi deteksi serangan web melalui kemampuan analisis pola secara otomatis. Temuan dari berbagai studi memperlihatkan bahwa algoritma berbasis *supervised learning* banyak digunakan karena mampu memanfaatkan data berlabel untuk mengidentifikasi anomali dengan lebih tepat [1], [4]. Analisis juga menunjukkan bahwa model *deep learning* memberikan performa lebih tinggi pada skenario serangan kompleks melalui kemampuan ekstraksi fitur yang lebih mendalam [11]. Indikasi ini memperlihatkan bahwa pendekatan komputasional adaptif menjadi elemen kunci pada mekanisme deteksi serangan web berbasis data.

Jika dibandingkan dengan penelitian terdahulu yang memanfaatkan metode tradisional seperti *signature-based detection*, hasil kajian ini menunjukkan bahwa model *machine learning* lebih efektif dalam mengenali variasi serangan baru [3]. Sejumlah penelitian terdahulu menekankan keterbatasan pendekatan konvensional dalam menghadapi pola serangan dinamis, sedangkan kajian ini menunjukkan bahwa model berbasis pembelajaran adaptif mampu mengatasi kendala tersebut. Konsistensi temuan antar penelitian menguatkan posisi algoritma modern sebagai pendekatan yang lebih fleksibel untuk deteksi ancaman. Kajian ini juga memperlihatkan bahwa integrasi berbagai teknik prapemrosesan data memengaruhi sensitivitas model dalam mengenali pola serangan halus.

Meskipun sebagian besar penelitian menunjukkan performa yang baik, terdapat beberapa hasil yang menunjukkan ketidakkonsistenan akurasi akibat ketidakseimbangan dataset atau kurangnya representasi pola serangan tertentu. Beberapa studi melaporkan bahwa model tertentu mengalami penurunan sensitivitas ketika data pelatihan mengandung bias distribusi [10]. Faktor ini dapat dijelaskan oleh keterbatasan model yang cenderung beradaptasi terhadap pola mayoritas dan mengabaikan variasi minor yang mewakili serangan baru. Fenomena ini menunjukkan pentingnya pengelolaan dataset untuk menjaga generalisasi model deteksi serangan web.

Secara teoritis, hasil kajian ini memperlihatkan bahwa penggunaan *machine learning* memperkaya fondasi ilmiah dalam studi keamanan jaringan melalui pemanfaatan pendekatan data-driven. Penggunaan algoritma klasifikasi, regresi, dan model *deep learning* memberikan gambaran menyeluruh tentang evolusi sistem deteksi serangan berbasis pembelajaran mesin [12]. Dari sisi praktis, temuan ini menunjukkan bahwa organisasi dapat meningkatkan efektivitas sistem keamanan dengan mengintegrasikan model adaptif yang mampu mempelajari karakteristik trafik web secara berkelanjutan. Keterbatasan kajian ini terletak pada variasi metodologis dan performa antar penelitian sehingga generalisasi hasil memerlukan kehati-hatian pada konteks implementasi tertentu.

6. KESIMPULAN DAN SARAN

6.1. *Kesimpulan*

Berdasarkan kajian sistematis terhadap 15 artikel internasional menunjukkan bahwa penerapan *Machine Learning (ML)* dan *Deep Learning (DL)* berperan penting dalam meningkatkan efektivitas *web attack detection*. Terdapat pergeseran signifikan dari algoritma klasik seperti *Random Forest (RF)* dan *Support Vector Machine (SVM)* menuju arsitektur *Deep Learning* dan *Transformer* yang lebih adaptif terhadap pola serangan kompleks seperti *Cross-Site Scripting (XSS)* dan *SQL Injection (SQLi)*. Secara umum, model-model tersebut mencapai performa tinggi (akurasi 85–97%, *F1-score* >90%), namun masih menghadapi kendala pada representativitas *dataset*, *class imbalance*, serta efisiensi komputasi. Studi ini menegaskan bahwa pemilihan algoritma deteksi harus disesuaikan dengan karakteristik data dan kebutuhan sistem. Integrasi *Explainable Artificial Intelligence (XAI)*, *Federated Learning (FL)*, dan *adversarial robustness* menjadi arah penting untuk mewujudkan sistem deteksi serangan web yang andal, efisien, dan transparan.

6.2. *Saran*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Penelitian selanjutnya disarankan untuk mengeksplorasi model ringan (*lightweight models*) berbasis *Transformer* dan *Federated Learning* agar dapat diterapkan pada perangkat *Internet of Things (IoT)* dengan sumber daya terbatas. Penggunaan *real-world datasets* yang lebih representatif serta penerapan *Explainable AI (XAI)* juga diperlukan guna meningkatkan interpretabilitas dan kepercayaan pengguna. Pendekatan berbasis *adversarial training* dan *defensive distillation* direkomendasikan untuk memperkuat ketahanan model terhadap serangan baru, sehingga sistem deteksi dapat diimplementasikan secara adaptif dan berkelanjutan di lingkungan siber modern. Selain itu, pengujian lintas lingkungan (cross-environment evaluation) dan studi komparatif antar model diharapkan dapat memberikan wawasan lebih dalam mengenai kestabilan performa model terhadap variasi pola serangan yang muncul di dunia nyata.

DAFTAR PUSTAKA

[1] R. Z. Muttaqin and D. Sudiana, "Design of Realtime Web Application Firewall on Deep Learning-Based to Improve Web Application Security," *J. Penelit. Pendidik. IPA*, vol. 10, no. 12, pp. 11121–11129, 2025, doi: 10.29303/jppipa.v10i12.8346.

[2] W. H. Aljuaid and S. S. Alshamrani, "A Deep Learning Approach for Intrusion Detection Systems in Cloud Computing Environments," *Appl. Sci.*, vol. 14, no. 13, p. 5381, 2024, doi: 10.3390/app14135381.

[3] N. W. Khan *et al.*, "A Hybrid Deep Learning-Based Intrusion Detection System for IoT Networks," *Math. Biosci. Eng.*, vol. 20, no. 8, pp. 13491–13520, 2023, doi: 10.3934/mbe.2023602.

[4] R. Chaganti, W. Suliman, V. Ravi, and A. Dua, "Deep Learning Approach for SDN-Enabled Intrusion Detection System in IoT Networks," *Inf.*, vol. 14, no. 1, pp. 1–21, 2023, doi: 10.3390/info14010041.

[5] R. A. Elsayed, R. A. Hamada, M. I. Abdalla, and S. A. Elsaid, "Securing IoT and SDN Systems Using Deep-Learning Based Automatic Intrusion Detection," *Ain Shams Eng. J.*, vol. 14, no. 10, p. 102211, 2023, doi: 10.1016/j.asej.2023.102211.

[6] N. Nikhlis, M. Jamal, and U. Ghofur, "Deteksi Dini Pencurian Data pada Perangkat Seluler Menggunakan Machine Learning," *J. Ilm. Sist. Inf.*, vol. 4, no. 2, pp. 110–129, May 2025, doi: 10.51903/fz0xzt26.

[7] A. M. Asmaraloka, K. Nisa, M. A. Hermansyah, F. H. D. Saputra, and A. Setiawan, "Implementation of the K-Nearest Neighbor (KNN) Algorithm in Handwritten Digit Pattern Recognition Using the Zoning Method," *J. Ilm. Sist. Inf.*, vol. 4, no. 2, pp. 175–185, May 2025, doi: 10.51903/kf6s5f56.

[8] F. Syaila, P. Azzahra, and C. Printo, "Tinjauan Literatur Sistematis: Analisis Implementasi Kecerdasan Buatan untuk Verifikasi Dokumen," *J. Ilm. Sist. Inf.*, vol. 4, no. 3, pp. 417–430, Aug. 2025, doi: 10.51903/kjjwk708.

[9] R. W. Shiddiq, N. Karna, and I. D. Irawati, "Optimizing Machine Learning-Based Network Intrusion Detection System with Oversampling, Feature Selection and Extraction," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 11, no. 2, pp. 225–237, 2025, doi: 10.26555/jiteki.v11i2.30675.

[10] N. Mohamed, "Current Trends in AI and ML for Cybersecurity: A State-of-the-Art Survey," *Cogent Eng.*, vol. 10, no. 2, p. 2272358, 2023, doi: 10.1080/23311916.2023.2272358.

[11] A. Salam, F. Ullah, F. Amin, and A. Mohammad, "Deep Learning Techniques for Web-Based Attack Detection in," *Mdpi*, vol. 11, no. 4, pp. 1–18, 2023, doi: 10.3390/technologies11040107.

[12] R. L. Alaoui and E. H. Nfaoui, "Deep Learning for Vulnerability and Attack Detection on Web Applications: A Systematic Literature Review," *Futur. Internet*, vol. 14, no. 4, p. 118, 2022, doi: 10.3390/fi14040118.

[13] T. Subramani, V. Jeganathan, and S. Kunkuma Balasubramanian, "Machine learning and deep learning techniques for poultry tasks management: a review," *Multimed. Tools Appl.*, vol. 84, no. 2, pp. 603–645, 2025, doi: 10.1007/s11042-024-18951-0.

[14] S. Sagar and C. Keke, "Confidential Machine Learning on Untrusted Platforms: A Survey," *Cybersecurity*, vol. 4, no. 1, p. 30, 2021, doi: 10.1186/s42400-021-00092-8.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[15] O. Arreche, T. Guntur, and M. Abdallah, "XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems," *Appl. Sci.*, vol. 14, no. 10, p. 4170, 2024, doi: 10.3390/app14104170.

[16] A. Alabbadi and F. Bajaber, "An Intrusion Detection System over the IoT Data Streams Using eXplainable Artificial Intelligence (XAI)," *Sensors*, vol. 25, no. 3, p. 847, 2025, doi: 10.3390/s25030847.

[17] E. M. Timofte *et al.*, "Federated Learning for Cybersecurity: A Privacy-Preserving Approach," *Appl. Sci.*, vol. 15, no. 12, pp. 1–27, 2025, doi: 10.3390/app15126878.

[18] M. Rahmati and A. Pagano, "Federated Learning-Driven Cybersecurity Framework for IoT Networks with Privacy Preserving and Real-Time Threat Detection Capabilities," *Informatics*, vol. 12, no. 3, p. 62, 2025, doi: 10.3390/informatics12030062.

[19] L. Segurola-Gil, M. Moreno-Moreno, I. Irigoien, and A. M. Florez-Tapia, "Unsupervised Anomaly Detection Approach for Cyberattack Identification," *Int. J. Mach. Learn. Cybern.*, vol. 15, no. 11, pp. 5291–5302, 2024, doi: 10.1007/s13042-024-02237-w.

[20] H. Park, D. Shin, C. Park, J. Jang, and D. Shin, "Unsupervised Machine Learning Methods for Anomaly Detection in Network Packets," *Electronics*, vol. 14, no. 14, p. 2779, 2025, doi: 10.3390/electronics14142779.

[21] R. T. Lo, W. J. Hwang, and T. M. Tai, "SQL Injection Detection Based on Lightweight Multi-Head Self-Attention," *Appl. Sci.*, vol. 15, no. 2, pp. 1–17, 2025, doi: 10.3390/app15020571.

[22] F. Younas, A. Raza, N. Thalji, L. Abualigah, R. A. Zitar, and H. Jia, "An Efficient Artificial Intelligence Approach for Early Detection of Cross-Site Scripting Attacks," *Decis. Anal. J.*, vol. 11, p. 100466, 2024, doi: 10.1016/j.dajour.2024.100466.

[23] R. Alhamyani and M. Alshammari, "Machine Learning-Driven Detection of Cross-Site Scripting Attacks," *Information*, vol. 15, no. 7, p. 420, 2024, doi: 10.3390/info15070420.

[24] M. AlJamal *et al.*, "A Robust Machine Learning Model for Detecting XSS Attacks on IoT over 5G Networks," *Futur. Internet*, vol. 16, no. 12, pp. 1–18, 2024, doi: 10.3390/fi16120482.

[25] R. Alqura'n *et al.*, "Advancing XSS Detection in IoT over 5G: A Cutting-Edge Artificial Neural Network Approach," *Internet of Things*, vol. 5, no. 3, pp. 478–508, 2024, doi: 10.3390/iot5030022.

[26] I. S. Crespo-Martínez, A. Campazas-Vega, Á. M. Guerrero-Higueras, V. Riego-DelCastillo, C. Álvarez-Aparicio, and C. Fernández-Llamas, "SQL Injection Attack Detection in Network Flow Data," *Comput. Secur.*, vol. 127, p. 103093, 2023, doi: 10.1016/j.cose.2023.103093.

[27] S. K. Erskine, "Real-Time Large-Scale Intrusion Detection and Prevention System (IDPS) CICIoT Dataset Traffic Assessment Based on Deep Learning," *Appl. Syst. Innov.*, vol. 8, no. 2, p. 52, 2025, doi: 10.3390/asi8020052.

[28] X. Li *et al.*, "An LSTM Based Cross-Site Scripting Attack Detection Scheme for Cloud Computing Environments," *J. Cloud Comput.*, vol. 12, no. 1, p. 118, 2023, doi: 10.1186/s13677-023-00483-x.

[29] A. H. Salem, S. M. Azzam, O. E. Emam, and A. A. Abohany, *Advancing Cybersecurity: A Comprehensive Review of AI-Driven Detection Techniques*, vol. 11, no. 1. Springer International Publishing, 2024. doi: 10.1186/s40537-024-00957-y.

[30] R. Vadisetty *et al.*, "Intelligent Detection of Injection Attacks via SQL Based on Supervised Machine Learning Models for Enhancing Web Security," *J. Artif. Intell. Big Data*, vol. 4, no. 2, pp. 109–119, 2024, doi: 10.31586/jaibd.2024.1333.

[31] D. P. Purbawa, A. J. Ulhaq, G. Ikhsan, and A. M. Shiddiqi, "An Enhanced SQL Injection Detection Using Ensemble Method," *JUTI J. Ilm. Teknol. Inf.*, vol. 21, no. 1, pp. 1–9, 2023, doi: 10.12962/j24068535.v21i1.a1060.

[32] M. K. Baklizi *et al.*, "Web Attack Intrusion Detection System Using Machine Learning Techniques,"

Int. J. online Biomed. Eng., vol. 20, no. 3, pp. 24–38, 2024, doi: 10.3991/ijoe.v20i03.45249.

[33] S. H. Rafique, A. Abdallah, N. S. Musa, and T. Murugan, "Machine Learning and Deep Learning Techniques for Internet of Things Network Anomaly Detection—Current Research Trends," *Sensors (Basel)*, vol. 24, no. 6, p. 1968, 2024, doi: 10.3390/s24061968.

[34] N. P. Bora, P. M. Bora, V. S. Tidake, G. P. Dhomse, and D. P. Pawar, "AI-Driven Biometric Captcha: Defending Against Automated Threats in Web Security," *Int. J. Basic Appl. Sci.*, vol. 14, no. 1, pp. 32–39, 2025, doi: 10.14419/fab5t109.

[35] W. Kanakri and B. King, "PK-Judge: Enhancing IP Protection of Neural Network Models Using an Asymmetric Approach," *Big Data Cogn. Comput.*, vol. 9, no. 3, p. 66, 2025, doi: 10.3390/bdcc9030066.

[36] B. Buyuktanir, Ş. Altinkaya, G. Karatas Baydogmus, and K. Yildiz, "Federated Learning in Intrusion Detection: Advancements, Applications, and Future Directions," *Cluster Comput.*, vol. 28, no. 7, p. 66, 2025, doi: 10.1007/s10586-025-05325-w.