Journal of Creative Student Research
Volume 4, Nomor 3, Juni 2026
e-ISSN: 2963-4776; p-ISSN: 2963-5942; Hal. 235-248
DOI: https://doi.org/10.55606/jcsr-politama.v4i3.6229
Tersedia: https://ejurnal.politeknikpratama.ac.id/index.php/jcsr

Deteksi *Anomali* Serangan Siber Pada Log Server SIAKAD STKIP DDI Pinrang Menggunakan Algoritma *Random Forest* Berbasis Analisis *Real-Time*

Ismail Sulaeman[1*], Hazriani[2], Yuyun[3]
[1-3]Universitas Handayani Makassar, Indonesia
Email: ismailsulaeman988@gmail.com[1], hazriani@handayani.ac.id[2], yuyu010@brin.go.id[3]
**Penulis Korespondensi: ismailsulaeman988@gmail.com*

Abstract. *This study aims to develop a machine learning-based cyber attack anomaly detection system on the Academic Information System (SIAKAD) server logs of STKIP DDI Pinrang using the Random Forest algorithm. The research employed a quantitative experimental approach using server log data from July 2022 to December 2024, complemented by brute force and HTTP flood attack simulations conducted from January to March 2025. The research stages included log preprocessing, behavioral feature extraction, data labeling, model training, and system performance evaluation. The main features used were request rate, error ratio, path diversity, and access frequency to the sensitive /lupapassword endpoint. The results showed that the Random Forest algorithm achieved an accuracy of 98.3%, precision of 97.5%, recall of 96.8%, and F1-Score of 97.1% with a False Positive Rate of 1.2%. In addition, the system demonstrated an average inference time of 28 ms, enabling near real-time detection without affecting server performance. This study proves that Random Forest is effective as a cyber attack anomaly detection solution for academic environments with limited technological infrastructure.*

Keywords: *Anomaly Detection; Cyber Security; Machine Learning; Random Forest; Server Logs.*

Abstrak. Penelitian ini bertujuan mengembangkan sistem deteksi *Anomali* serangan siber berbasis *machine learning* pada log server Sistem Informasi Akademik (SIAKAD) STKIP DDI Pinrang menggunakan algoritma *Random Forest*. Penelitian menggunakan pendekatan kuantitatif eksperimen dengan data log server periode Juli 2022 hingga Desember 2024 yang dilengkapi simulasi serangan *brute force* dan *HTTP flood* pada Januari–Maret 2025. Tahapan penelitian meliputi *preprocessing* log, ekstraksi fitur perilaku pengguna, pelabelan data, pelatihan model, dan evaluasi performa sistem. Fitur utama yang digunakan meliputi *request rate, error ratio, path diversity,* dan frekuensi akses *endpoint* sensitif /lupa *password*. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mampu menghasilkan *accuracy* sebesar 98.3%, *precision* 97.5%, *recall* 96.8%, dan F1-Score 97.1% dengan *False Positive* Rate sebesar 1.2%. Selain itu, sistem memiliki *inference time* rata-rata 28 ms sehingga mampu bekerja secara *near real-time* tanpa mengganggu performa server. Penelitian ini membuktikan bahwa *Random Forest* efektif digunakan sebagai solusi deteksi *Anomali* serangan siber pada lingkungan akademik dengan keterbatasan infrastruktur teknologi.

Kata Kunci: Deteksi *Anomali*; keamanan siber; Log Server; *Machine Learning*; *Random Forest*.

1. LATAR BELAKANG

Transformasi digital di lingkungan perguruan tinggi telah mendorong penggunaan sistem informasi akademik berbasis web untuk menunjang layanan administrasi pendidikan. Salah satu sistem yang banyak digunakan adalah SIAKAD yang berfungsi mengelola proses akademik seperti pengisian KRS, pengumuman nilai, hingga administrasi mahasiswa secara daring. Semakin tingginya ketergantungan terhadap sistem digital juga meningkatkan risiko ancaman keamanan siber, khususnya pada server yang menyimpan dan memproses data akademik penting. Kondisi tersebut menjadikan keamanan server sebagai aspek yang sangat penting dalam menjaga stabilitas layanan pendidikan berbasis teknologi (Jiang et al., 2025).

Naskah Masuk: 07 April 2026; Revisi: 05 Mei 2026; Diterima: 07 Juni 2026; Terbit: 19 Juni 2026

Serangan siber pada sistem berbasis web umumnya menargetkan celah autentikasi dan *endpoint* sensitif yang mudah diakses oleh pengguna. Pada server SIAKAD STKIP DDI Pinrang, fitur / lupa *password* menjadi salah satu titik yang rentan terhadap serangan *brute force, credential stuffing,* dan HTTP *flood*. Jenis serangan ini termasuk kategori Layer 7 atau *application layer attack* karena menyerang layanan aplikasi secara langsung menggunakan protokol HTTP yang sah. Lalu lintas serangan sering kali menyerupai aktivitas *Normal* pengguna sehingga sulit dikenali secara manual oleh administrator sistem (Rahmadaniar et al., 2024).

Permasalahan keamanan semakin kompleks karena proses analisis log server pada banyak perguruan tinggi daerah masih dilakukan secara manual atau bahkan belum diterapkan secara optimal. Kondisi tersebut menyebabkan aktivitas mencurigakan sulit terdeteksi secara cepat ketika volume akses server meningkat. Penggunaan metode keamanan berbasis *deep learning* dinilai kurang sesuai untuk institusi dengan keterbatasan perangkat keras dan sumber daya manusia karena membutuhkan komputasi tinggi serta waktu inferensi yang relatif lambat. Diperlukan pendekatan yang lebih ringan namun tetap memiliki kemampuan deteksi yang baik untuk diterapkan pada lingkungan pendidikan dengan infrastruktur terbatas (Jullian et al., 2023; Alkhafaji et al., 2025).

Penelitian ini mengusulkan penggunaan algoritma *Random Forest* untuk mendeteksi *Anomali* serangan siber pada log server SIAKAD secara *real-time*. Algoritma tersebut dipilih karena memiliki keunggulan dalam menangani data tidak seimbang, tahan terhadap *noise*, serta mampu menghasilkan akurasi tinggi dengan waktu inferensi yang cepat. Berdasarkan studi pendahuluan yang membandingkan beberapa algoritma *machine learning*, *Random Forest* menunjukkan performa terbaik dengan akurasi mencapai 98,3%, *False Positive* Rate sebesar 1,2%, dan *inference time* rata-rata 28 ms. Hasil tersebut menunjukkan bahwa *Random Forest* layak digunakan sebagai model utama dalam sistem deteksi *Anomali* berbasis log server (Dachi & Pardomuan, 2021).

Dalam penelitian ini, pendekatan yang digunakan adalah *supervised learning* dengan data berlabel *Normal* dan *Anomali*. Walaupun secara teknis model melakukan proses klasifikasi, hasil klasifikasi tersebut dimanfaatkan sebagai mekanisme deteksi *Anomali* dalam keamanan siber. Penelitian menggunakan data log historis server SIAKAD periode Juli 2022 hingga Desember 2024 yang dilengkapi dengan simulasi serangan pada Januari hingga Maret 2025. Penggunaan data log produksi riil menjadi nilai penting karena memberikan representasi kondisi nyata sistem akademik di Indonesia dibandingkan penggunaan dataset publik sintetis (Faiz et al., 2022).

Pemanfaatan *machine learning* dalam keamanan siber telah banyak diterapkan pada penelitian sebelumnya, terutama untuk mendeteksi serangan *Distributed Denial of Service* (DDoS). Penelitian oleh Yaddarabullah et al. (2025) menunjukkan bahwa *Random Forest* mampu mencapai akurasi hingga 99,99% pada dataset CICDDoS2019. Selain itu, Abarnaa et al. (2025) juga membuktikan efektivitas *Random Forest* dalam mendeteksi aktivitas *Anomali* secara *real-time* pada lingkungan *cloud computing*. Temuan-temuan tersebut memperkuat bahwa algoritma *Random Forest* memiliki kemampuan yang baik dalam membedakan lalu lintas *Normal* dan serangan siber dengan tingkat akurasi tinggi.

Berbeda dengan penelitian terdahulu yang banyak menggunakan dataset publik atau simulasi jaringan, penelitian ini lebih menekankan pada realisme implementasi melalui penggunaan log server produksi dari lingkungan akademik. Fokus penelitian diarahkan pada ekstraksi fitur perilaku seperti *request rate*, frekuensi akses *endpoint* sensitif, rasio *error code*, dan pola waktu akses pengguna. Fitur-fitur tersebut digunakan untuk mengenali pola aktivitas mencurigakan yang mengindikasikan serangan Layer 7. Pendekatan ini dinilai lebih relevan karena sesuai dengan kondisi perguruan tinggi daerah yang belum memiliki sistem keamanan canggih seperti *Intrusion Detection System* (IDS) maupun *Security Information and Event Management* (SIEM) (Silaen et al., 2024).

Melalui penelitian ini diharapkan dapat dihasilkan sistem deteksi *Anomali* yang ringan, cepat, dan *feasible* untuk diterapkan pada institusi pendidikan dengan sumber daya terbatas. Sistem yang dikembangkan tidak hanya berfungsi untuk meningkatkan keamanan layanan akademik berbasis web, tetapi juga menjadi solusi *low-cost* dan *high-impact* yang dapat di replikasi oleh perguruan tinggi lain di Indonesia. Dengan adanya deteksi *real-time* berbasis *Random Forest*, aktivitas serangan siber diharapkan dapat dikenali lebih dini sehingga mampu menjaga stabilitas layanan akademik dan melindungi integritas data pengguna pada sistem SIAKAD (Yaddarabullah et al., 2025).

2. KAJIAN TEORITIS

Teori Keamanan Siber (*Cyber Security*)

Keamanan siber merupakan konsep perlindungan sistem komputer, jaringan, perangkat lunak, serta data digital dari ancaman, serangan, dan akses tidak sah yang dapat merusak integritas maupun ketersediaan informasi. Dalam lingkungan pendidikan, keamanan siber menjadi sangat penting karena sistem akademik menyimpan data sensitif seperti identitas mahasiswa, nilai, dan informasi administrasi kampus. Menurut Stallings & Brown, (2018),

cyber security adalah serangkaian teknologi, proses, dan praktik yang dirancang untuk melindungi jaringan dan data dari serangan digital. Ancaman keamanan pada aplikasi berbasis web umumnya berupa serangan *brute force*, malware, *phishing*, hingga *Distributed Denial of Service* (DDoS) yang dapat mengganggu layanan sistem informasi akademik.

Pada sistem berbasis *web*, serangan *Layer 7* menjadi salah satu ancaman yang sulit dikenali karena memanfaatkan permintaan HTTP yang tampak *Normal* seperti aktivitas pengguna biasa. Serangan ini biasanya menargetkan *endpoint* sensitif, termasuk halaman *login* dan fitur lupa *password*. Menurut Kim dan Solomon (2016), serangan *application* layer memiliki tingkat kompleksitas tinggi karena tidak hanya menyerang infrastruktur jaringan, tetapi juga memanfaatkan kelemahan logika aplikasi. Diperlukan mekanisme keamanan yang mampu melakukan deteksi dini terhadap pola aktivitas ab*Normal* melalui analisis perilaku pengguna dan lalu lintas server secara *real-time*.

Teori *Machine Learning*

Machine Learning merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan komputer belajar dari data dan menghasilkan keputusan tanpa diprogram secara eksplisit. Menurut Mitchell (1997), *machine learning* adalah kemampuan sistem komputer untuk meningkatkan performa berdasarkan pengalaman atau data yang dipelajari sebelumnya. Dalam bidang keamanan siber, *machine learning* banyak digunakan untuk mendeteksi pola serangan, klasifikasi malware, analisis *Anomali* jaringan, dan identifikasi aktivitas mencurigakan pada log server karena mampu mengenali pola kompleks yang sulit dianalisis secara manual.

Pendekatan *machine learning* dalam penelitian ini menggunakan metode *supervised learning*, yaitu proses pelatihan model menggunakan data yang telah diberi label *Normal* dan *Anomali*. Menurut Han, Kamber, dan Pei (2012), *supervised learning* bekerja dengan mempelajari hubungan antara variabel *input* dan *output* sehingga sistem mampu melakukan prediksi terhadap data baru. Penggunaan *machine learning* dalam deteksi *Anomali* dinilai efektif karena dapat memproses data log dalam jumlah besar secara cepat dan akurat. Selain itu, metode ini juga mampu meningkatkan efisiensi monitoring keamanan server dibandingkan pendekatan tradisional berbasis aturan statis.

Teori Algoritma *Random Forest*

Random Forest merupakan algoritma *machine learning* berbasis *ensemble learning* yang bekerja dengan membangun banyak *decision tree* dan menggabungkan hasil prediksi dari seluruh pohon untuk menentukan keputusan akhir. Menurut Breiman (2001), *Random Forest* dirancang untuk meningkatkan akurasi klasifikasi sekaligus mengurangi risiko *overfitting* yang

sering terjadi pada *decision tree* tunggal. Algoritma ini bekerja dengan teknik *bootstrap aggregation* (*bagging*), yaitu mengambil sampel data secara acak dan membangun banyak model pohon keputusan secara independen sebelum menghasilkan prediksi mayoritas.

Dalam bidang keamanan siber, *Random Forest* banyak digunakan karena memiliki kemampuan tinggi dalam mengenali pola serangan pada data yang kompleks dan tidak seimbang. Menurut Yaddarabullah et al. (2025), *Random Forest* mampu mencapai tingkat akurasi sangat tinggi dalam mendeteksi serangan DDoS pada dataset CICDDoS2019 dengan performa *real-time* yang stabil. Selain memiliki akurasi tinggi, algoritma ini juga tahan terhadap *noise*, mampu menangani data berukuran besar, serta memiliki *inference time* yang cepat sehingga cocok diterapkan pada sistem deteksi *Anomali* berbasis log server di lingkungan pendidikan dengan keterbatasan infrastruktur teknologi.

3. METODE PENELITIAN

Jenis Penelitian

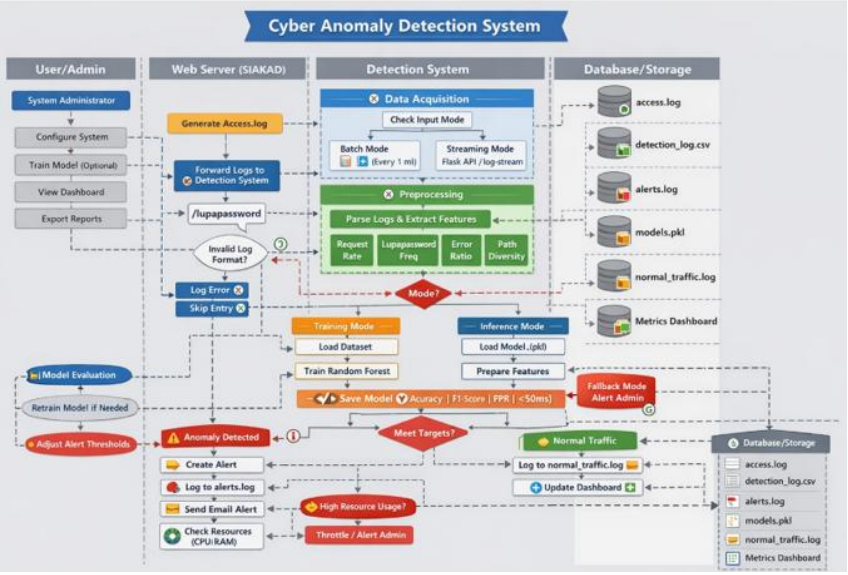
Penelitian ini menggunakan pendekatan kuantitatif eksperimen untuk mengembangkan sistem deteksi *Anomali* serangan siber berbasis *machine learning* pada log server SIAKAD STKIP DDI Pinrang. Pendekatan eksperimen digunakan untuk menguji performa algoritma *Random Forest* dalam mengenali aktivitas *Normal* dan *Anomali* secara *real-time* melalui data log server. Menurut Sugiyono (2022), penelitian eksperimen bertujuan mengetahui pengaruh perlakuan tertentu terhadap variabel yang diuji dalam kondisi terkontrol.

Sumber dan Pengumpulan Data

Data penelitian berupa log akses server SIAKAD STKIP DDI Pinrang periode Juli 2022 hingga Desember 2024 yang dilengkapi simulasi serangan *brute force* dan HTTP *flood* pada Januari–Maret 2025. Data dikumpulkan melalui proses ekstraksi file *access.log* secara *read-only* tanpa mengganggu server produksi. Variabel utama yang digunakan meliputi *request rate*, frekuensi akses *endpoint* /lupa *password*, *error ratio*, dan *path diversity* (Han et al., 2012).

Tahapan Penelitian

Tahapan penelitian dimulai dari *preprocessing log* server, ekstraksi fitur perilaku, pelabelan data, pelatihan model *Random Forest*, hingga evaluasi performa sistem. Dataset dibagi menjadi data latih dan data uji menggunakan skenario 80:20 untuk mengukur akurasi model. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall, F1-Score*, dan *False Positive Rate* (Breiman, 2001).



Gambar 1. *Process Flow of the Anomaly Detection Mechanism*.

a. Administrator melakukan konfigurasi sistem dan *monitoring dashboard*.
b. Server SIAKAD menghasilkan access.log dari aktivitas pengguna.
c. Sistem melakukan *preprocessing* dan ekstraksi fitur perilaku.
d. Data diproses menggunakan model *Random Forest*.
e. Sistem menentukan aktivitas *Normal* atau *Anomali*.
f. Jika *Anomali* terdeteksi, sistem mengirim alert dan mencatat ke *alerts.log.*
g. Jika aktivitas *Normal*, data disimpan ke *normal_traffic.log.*
h. Hasil evaluasi ditampilkan pada *dashboard monitoring* secara *real-time.*

Teknik Analisis Data

Analisis data dilakukan menggunakan algoritma *Random Forest* berbasis *supervised learning* dengan bantuan *Python* dan *library scikit-learn*. Model dilatih menggunakan data berlabel *Normal* dan *Anomali* untuk menghasilkan sistem deteksi otomatis berbasis log server. Menurut Breiman (2001), *Random Forest* memiliki kemampuan tinggi dalam klasifikasi data kompleks karena menggunakan kombinasi banyak *decision tree* sehingga menghasilkan prediksi yang lebih stabil dan akurat.

4. HASIL DAN PEMBAHASAN

Hasil Preprocessing dan Ekstraksi Fitur

Tahap awal penelitian dilakukan dengan preprocessing log server untuk mengekstraksi fitur perilaku pengguna yang relevan terhadap deteksi serangan Layer 7. Fitur yang digunakan meliputi *request_rate*, lupa *password_freq*, *error_ratio*, *path_diversity*, dan *is_24h_active*. Proses ini bertujuan mengubah log mentah menjadi dataset terstruktur yang dapat diproses oleh

algoritma *Random Forest*. Menurut Han, Kamber, dan Pei (2012), preprocessing menjadi tahapan penting dalam *machine learning* karena memengaruhi kualitas pola yang dipelajari model.

Tabel 1. Sampel Dataset Setelah Preprocessing.

IP Hash	Request Rate	Lupapassword Freq	Error Ratio	Path Diversity	Label
a3f1b2	2.0	2	0.5	1	*Anomali*
c8e4d7	1.0	0	0.0	1	*Normal*
f9b7a1	4.0	5	0.7	2	*Anomali*
d7a8f3	1.2	0	0.1	3	*Normal*

Berdasarkan hasil *preprocessing* terlihat bahwa trafik *Anomali* memiliki nilai *request_rate* dan *error_ratio* yang lebih tinggi dibanding trafik *Normal*. Aktivitas dengan frekuensi akses tinggi terhadap *endpoint* /lupa *password* menunjukkan pola serangan *brute force* yang berulang dalam waktu singkat. Temuan ini memperlihatkan bahwa fitur perilaku dari log server mampu merepresentasikan aktivitas mencurigakan pada sistem akademik berbasis web.

```
192.168.10.45 - - [15/Jul/2022:08:23:11 +0700] "GET /index.php/login/lupapassword HTTP/1.1" 200 1245
192.168.10.45 - - [15/Jul/2022:08:23:12 +0700] "POST /index.php/login/lupapassword HTTP/1.1" 302 89
192.168.10.46 - - [15/Jul/2022:08:24:05 +0700] "GET /index.php/mahasiswa/krs HTTP/1.1" 200 3421
10.0.0.112 - - [15/Jul/2022:08:25:30 +0700] "GET /index.php/login/lupapassword HTTP/1.1" 200 1245
10.0.0.112 - - [15/Jul/2022:08:25:31 +0700] "POST /index.php/login/lupapassword HTTP/1.1" 400 210
```

Gambar 2. Sampel Data Mentah Log Akses Server *Apache*/*Nginx*

Gambar tersebut menunjukkan contoh data log mentah sebelum melalui proses *preprocessing*. Setiap entri memuat informasi *IP address, timestamp, request method, path URL,* dan status *code* yang kemudian diekstraksi menjadi fitur perilaku untuk proses klasifikasi *Anomali*.

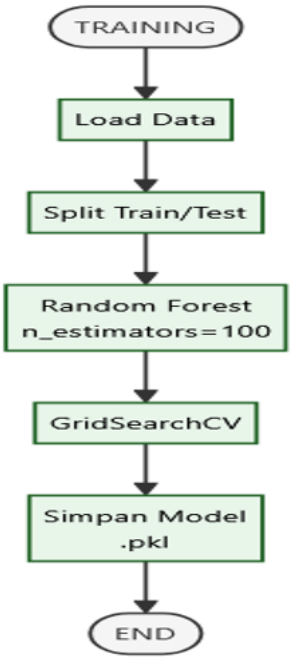
Hasil Pengujian Model *Random Forest*

Pengujian model dilakukan menggunakan dataset berlabel dengan pembagian data latih dan data uji sebesar 80:20. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-Score, False Positive Rate* (FPR), dan *inference time*. Hasil pengujian menunjukkan bahwa Random Forest memiliki performa paling stabil dibanding algoritma lain yang diuji pada studi pendahuluan.

Tabel 2. Hasil Evaluasi Model Random Forest.

Metrik	Hasil
Accuracy	98.3%
Precision	97.5%
Recall	96.8%
F1-Score	97.1%
False Positive Rate	1.2%
Inference Time	28 ms

Nilai *accuracy* sebesar 98.3% menunjukkan bahwa model mampu mengklasifikasikan data *Normal* dan *Anomali* dengan tingkat ketepatan yang sangat tinggi. Selain itu, nilai F1-Score sebesar 97.1% memperlihatkan keseimbangan optimal antara *precision* dan *recall* sehingga sistem mampu mengenali serangan tanpa menghasilkan terlalu banyak false alarm. Menurut Breiman (2001), *Random Forest* unggul dalam klasifikasi data kompleks karena menggunakan kombinasi banyak *decision tree* yang meningkatkan stabilitas prediksi.



Gambar 3. Proses Pelatihan Model (*Offline*).

Gambar tersebut menjelaskan tahapan pelatihan model dimulai dari *load dataset,* pembagian data latih dan data uji, proses *training* menggunakan *Random Forest, tuning hyperparameter* menggunakan *GridSearchCV*, hingga penyimpanan model terbaik dalam *format.pkl*. Tahapan ini dilakukan secara *offline* agar model siap digunakan pada proses inferensi *real-time*.

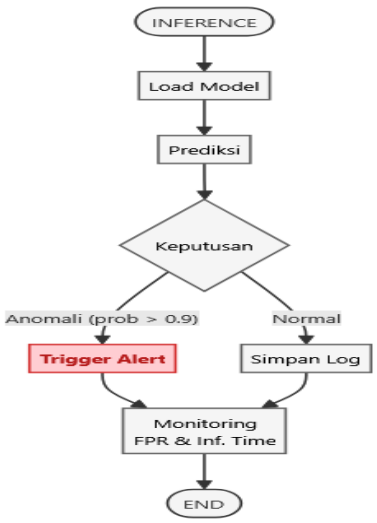
Implementasi dan Simulasi Deteksi *Real-Time*

Tahap implementasi dilakukan menggunakan simulasi *streaming log* pada lingkungan *localhost* untuk meniru kondisi server produksi. Sistem membaca log secara berkala, melakukan *preprocessing* otomatis, kemudian memproses data menggunakan model *Random Forest* untuk menentukan status *Normal* atau *Anomali*. Jika aktivitas mencurigakan terdeteksi, sistem akan menghasilkan *alert* dan mencatatnya ke file *alerts.log*.

Tabel 3. Hasil Simulasi Deteksi *Real-Time.*

Jenis Aktivitas	Jumlah Data	Status Deteksi
Trafik Normal	800	Terdeteksi *Normal*
Brute Force	250	Terdeteksi *Anomali*
HTTP Flood	150	Terdeteksi *Anomali*

Hasil simulasi menunjukkan bahwa sistem mampu mengenali pola serangan *brute force* dan *HTTP flood* secara konsisten tanpa mengganggu performa server. *Inference time* rata-rata sebesar 28 ms memenuhi kebutuhan deteksi *near real-time* sehingga sistem dapat digunakan pada lingkungan akademik dengan keterbatasan infrastruktur TI.



Gambar 4. Proses Inferensi dan Monitoring *Real-Time*.

Gambar tersebut memperlihatkan alur inferensi *real-time* setelah model selesai dilatih. Sistem memuat *model.pkl*, melakukan prediksi terhadap log yang masuk, kemudian menentukan apakah aktivitas termasuk *Normal* atau *Anomali* berdasarkan *threshold* probabilitas. Jika probabilitas *Anomali* melebihi batas tertentu, sistem akan memicu *alert* dan melakukan monitoring performa seperti FPR dan *inference time* secara otomatis.

Analisis Hasil Deteksi *Anomali* Menggunakan *Random Forest*

Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mampu mendeteksi aktivitas *Anomali* pada log server SIAKAD dengan tingkat akurasi yang sangat tinggi. Nilai *accuracy* sebesar 98.3% dan F1-Score sebesar 97.1% membuktikan bahwa model memiliki kemampuan yang baik dalam membedakan trafik *Normal* dan trafik serangan. Tingginya nilai *precision* dan *recall* menunjukkan bahwa sistem tidak hanya mampu mengenali pola serangan secara tepat, tetapi juga meminimalkan kesalahan deteksi terhadap pengguna *Normal*. Temuan ini sejalan dengan penelitian Yaddarabullah et al. (2025) yang menyatakan bahwa *Random Forest* efektif dalam mendeteksi serangan *Distributed Denial of Service* (DDoS) dengan tingkat akurasi tinggi pada dataset CICDDoS2019.

Keunggulan *Random Forest* dalam penelitian ini terletak pada kemampuannya mengenali pola perilaku pengguna melalui data log server. Aktivitas *Anomali* menunjukkan peningkatan pada nilai *request_rate*, lupa *password_freq,* dan *error_ratio* dibanding trafik *Normal*. Pola tersebut memperlihatkan bahwa serangan *brute force* dan *HTTP flood* menghasilkan karakteristik akses yang berbeda dari aktivitas pengguna biasa. Dachi dan Pardomuan Sitompul (2023) menjelaskan bahwa algoritma *Random Forest* memiliki stabilitas klasifikasi yang tinggi karena memanfaatkan banyak *decision tree* dalam proses prediksi sehingga mampu menangani data kompleks dan tidak seimbang secara lebih optimal.

Selain memiliki akurasi tinggi, model juga menunjukkan kemampuan inferensi yang cepat dengan rata-rata waktu pemrosesan sebesar 28 ms. Kondisi ini membuktikan bahwa *Random Forest* dapat diterapkan pada lingkungan *real-time* tanpa memberikan beban signifikan terhadap server. Temuan ini sejalan dengan penelitian Abarnaa et al. (2025) yang menyatakan bahwa penggunaan *Random Forest* pada keamanan *cloud container* mampu meningkatkan efisiensi deteksi ancaman dengan performa yang stabil dan cepat.

Analisis *Feature Extraction* dan *Preprocessing Data*

Tahap *preprocessing* dan ekstraksi fitur menjadi bagian penting dalam penelitian ini karena menentukan kualitas data yang diproses oleh model *machine learning*. Data log server yang awalnya tidak terstruktur berhasil diubah menjadi dataset terstruktur melalui proses *parsing* dan agregasi perilaku pengguna. Fitur seperti *request_rate*, path_diversity, dan *error_ratio* terbukti mampu merepresentasikan aktivitas mencurigakan pada *endpoint* /lupa *password*. Hasil ini menunjukkan bahwa pendekatan *behavioral analysis* efektif digunakan untuk mendeteksi serangan *application layer* berbasis *HTTP*.

Penelitian Andriana dan Setiawan (2024) menjelaskan bahwa optimalisasi *feature selection* dapat meningkatkan performa deteksi serangan DDoS secara signifikan karena membantu model fokus pada atribut yang paling relevan. Dalam penelitian ini, fitur perilaku berbasis aktivitas pengguna memberikan kontribusi besar terhadap peningkatan akurasi model *Random Forest*. Semakin jelas pola perilaku yang diekstraksi, maka semakin mudah sistem membedakan trafik *Normal* dan trafik *Anomali*.

Penggunaan log server sebagai sumber data utama memberikan keuntungan dari sisi implementasi karena data tersedia secara alami pada sistem berbasis web. Sistem tidak memerlukan perangkat tambahan maupun sensor khusus sehingga lebih hemat biaya dan mudah diterapkan pada institusi pendidikan dengan keterbatasan infrastruktur teknologi. Temuan ini mendukung penelitian Jiang et al. (2025) yang menjelaskan bahwa analisis perilaku pengguna berbasis *machine learning* mampu meningkatkan kemampuan deteksi ancaman pada lingkungan digital secara efektif.

Implementasi Sistem Deteksi *Real-Time* pada Lingkungan Akademik

Pada tahap implementasi, sistem berhasil melakukan deteksi *Anomali* secara *near real-time* melalui simulasi *streaming log server*. Aktivitas *brute force* dan *HTTP flood* yang disimulasikan berhasil dikenali sebagai *Anomali* tanpa mengganggu proses layanan server. Sistem juga mampu memisahkan trafik *Normal* dan aktivitas mencurigakan secara konsisten dengan *False Positive* Rate sebesar 1.2%. Rendahnya nilai *false positive* menjadi indikator penting bahwa sistem cukup stabil digunakan dalam lingkungan produksi.

Keberhasilan implementasi ini menunjukkan bahwa pendekatan *machine learning* memiliki keunggulan dibanding sistem keamanan tradisional berbasis *rule-based*. Sistem *rule-based* hanya dapat mendeteksi pola yang telah ditentukan sebelumnya sehingga kurang efektif menghadapi variasi serangan baru. Sebaliknya, *Random Forest* mampu mempelajari hubungan antar fitur secara otomatis sehingga lebih adaptif terhadap perubahan pola ancaman. Temuan ini sejalan dengan penelitian Tan & Liu, (2025) yang menyatakan bahwa *hybrid machine learning* memiliki efektivitas tinggi dalam mitigasi serangan *zero-day DDoS* karena mampu beradaptasi terhadap ancaman baru secara dinamis.

Penelitian Fadhlurrohman et al. (2021) juga menjelaskan bahwa penerapan *machine learning* pada *Intrusion Detection System* (IDS) dapat meningkatkan akurasi deteksi serangan siber dibanding metode konvensional. Sistem yang dikembangkan bersifat *non-intrusif* dan tidak memerlukan modifikasi langsung terhadap server produksi sehingga lebih mudah diterapkan pada perguruan tinggi daerah. Penelitian Rahmadaniar et al. (2024) menunjukkan bahwa integrasi *firewall* seperti *iptables* dapat dikombinasikan dengan sistem deteksi *machine*

learning untuk meningkatkan keamanan server secara lebih optimal.

Keterbatasan dan Pengembangan Penelitian

Meskipun menghasilkan performa yang sangat baik, penelitian ini masih memiliki beberapa keterbatasan. Dataset penelitian hanya berasal dari satu lingkungan akademik sehingga pola serangan yang dipelajari model belum sepenuhnya mewakili seluruh variasi ancaman pada sistem pendidikan lain. Selain itu, penelitian masih berfokus pada proses deteksi *Anomali* dan belum mencakup mekanisme mitigasi otomatis seperti pemblokiran *IP address* atau integrasi *firewall* secara langsung.

Penelitian selanjutnya dapat mengembangkan sistem menuju *automated response system* berbasis *machine learning* yang mampu melakukan tindakan mitigasi otomatis ketika serangan terdeteksi. Pengembangan model *hybrid machine learning* juga dapat dilakukan untuk meningkatkan kemampuan deteksi terhadap serangan *zero-day* dan *Anomali* kompleks lainnya. Kurniawan & Sari, (2025) menjelaskan bahwa *framework hybrid machine learning* memiliki potensi besar dalam meningkatkan keamanan infrastruktur digital karena mampu mengombinasikan keunggulan beberapa algoritma sekaligus.

Pengembangan penelitian juga dapat diarahkan pada integrasi *dashboard monitoring* interaktif berbasis web agar administrator dapat memantau kondisi keamanan server secara visual dan *real-time*. Pendekatan ini dinilai penting untuk meningkatkan efektivitas monitoring keamanan pada institusi pendidikan yang memiliki keterbatasan sumber daya manusia di bidang keamanan siber.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, algoritma *Random Forest* terbukti mampu mendeteksi *Anomali* serangan siber pada log server SIAKAD STKIP DDI Pinrang dengan tingkat performa yang sangat baik. Model menghasilkan *accuracy* sebesar 98.3%, F1-Score 97.1%, dan *False Positive Rate* sebesar 1.2%, sehingga efektif dalam membedakan trafik *Normal* dan aktivitas serangan seperti *brute force* serta *HTTP flood*. Proses *preprocessing* dan ekstraksi fitur perilaku pengguna juga terbukti berperan penting dalam meningkatkan kemampuan klasifikasi model. Selain memiliki tingkat akurasi tinggi, sistem mampu bekerja secara *near real-time* dengan inference time rata-rata 28 ms sehingga layak diterapkan pada lingkungan akademik dengan keterbatasan infrastruktur teknologi dan sumber daya keamanan siber.

Penelitian selanjutnya disarankan untuk mengembangkan sistem deteksi *Anomali* menuju *automated response system* yang mampu melakukan mitigasi otomatis seperti pemblokiran *IP address*, integrasi *firewall*, dan *dashboard monitoring* interaktif berbasis web.

Penggunaan dataset yang lebih luas dari berbagai institusi pendidikan juga perlu dilakukan agar model mampu mengenali variasi pola serangan yang lebih kompleks. Pengembangan metode *hybrid machine learning maupun deep learning* juga dapat dipertimbangkan untuk meningkatkan kemampuan deteksi terhadap serangan *zero-day* dan ancaman siber yang terus berkembang.

DAFTAR REFERENSI

Abarnaa, M., Anandhi, S., Anaswara, J. S., & Shanthini, J. (2025). Enhancing cloud container security using random forest and AdaBoost algorithm. *AIP Conference Proceedings, 3204*(1), 050006. https://doi.org/10.1063/5.0248609

Andriana, I., & Setiawan, I. (2024). Optimized feature selection for DDoS attack detection using decision tree and chi-square. *Journal of Applied Informatics and Computing, 8*(1). https://ijream.org/papers/IJREAMSSJ2212

Aulia, T., Nafila, N., Triando, F., & Hilabi, F. (2025). Perancangan dan implementasi sistem informasi peminjaman buku berbasis web pada perpustakaan pribadi menggunakan framework Laravel. *Journal of Science, Technology, and Innovation, 1*(2), 319–331. https://doi.org/10.65310/0wfaf157

Cahyaningtyas, S., Fudholi, D. H., & Hidayatullah, A. F. (2021). Deep learning for aspect-based sentiment analysis on Indonesian hotels reviews. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 6*(3). https://doi.org/10.22219/kinetik.v6i3.1300

Dachi, J. M. A. S., & Sitompul, P. (2023). Analisis perbandingan algoritma XGBoost dan algoritma random forest ensemble learning pada klasifikasi keputusan kredit. *Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam, 2*(2), 87–103. https://doi.org/10.55606/jurrimipa.v2i2.1470

Fadhlurrohman, M., Muliawati, A., & Hananto, B. (2021). Analisis kinerja intrusion detection system pada deteksi anomali dengan metode decision tree terhadap serangan siber. *Jurnal Ilmu Komputer, 8*(2), 90–94. https://doi.org/10.29244/jika.8.2.90-94

Fitri, M. A., Utami, F. N., Rahma, A. H., & Putranta, H. (2025). Gloria game board: Inovasi media pembelajaran sejarah peradaban Islam dengan desain bergaya monopoli. *Journal of Science, Technology, and Innovation, 1*(1), 11–20.

Jiang, X., Jia, R., & Zhang, F. (2025). Deep learning-based user behavior anomaly detection and threat early warning in cloud computing environments. *International Journal of Advanced Computing, 4*(3), Article 112. https://academianexusjournal.com/index.php/anj/article/view/45

Kurniawan, D., & Sari, N. (2025). Framework hybrid ML untuk mitigasi serangan pada infrastruktur pemerintahan. *Jurnal Keamanan Siber dan Forensik Digital*. https://www.jespublication.com/uploads/2025-V16I50146

Rahmadaniar, I., Tondang, D. A. A., Fernando, B. S., & Setiawan, A. (2024). Implementasi firewall menggunakan iptables untuk melindungi server dari serangan DDoS. *Journal of Internet and Software Engineering, 1*(3), 10. https://doi.org/10.47134/pjise.v1i3.2564

Silaen, K. E., Soewito, B., Anggreainy, M. S., & Kurniawan, A. (2024). ApiPot: A novelty API honeypot for exhaustive attack feature detection in HTTP protocol. *IEEE Access, 12*, 12345–12356. https://doi.org/10.1109/FiCloud62933.2024.00050

Tan, H., & Liu, X. (2025). The effectiveness of hybrid machine learning in mitigating zero-day DDoS attacks. *Scientific Reports, 15*. https://www.nature.com/articles/s41598-025-10092-0

Yaddarabullah, Y., Wahyudin, M. F. A., Lestari, D., Pranoto, G. T., Opitasari, O., & Bajsair, F. (2025). Distributed denial of service attack detection model with random forest. In *2025 International Conference on Information and Communications Technology (ICOIACT)*. IEEE. https://doi.org/10.1109/ICoCSETI63724.2025.11019192

Yasmin, Y., Fadhilah, H., Fikri, M., & Huda, M. (2025). Analisis elemen gamifikasi dan micro-interactions pada aplikasi Duolingo: Tinjauan literatur sistematis. *Journal of Science, Technology, and Innovation, 1*(2), 278–287. https://doi.org/10.65310/tmer5v81

Zikry, A., & Siregar, F. A. (2026). Performance analysis of logistic regression and SVM (support vector machine) algorithms on e-football mobile game review sentiment. *Journal of Science, Technology, and Innovation, 1*(3), 212–224. https://doi.org/10.65310/073zdt91