Sienna (E -ISSN: 2745-9861 | P-ISSN: 2745-987X)
https://jurnal.umko.ac.id/index.php/sienna
DOI: https://doi.org/10.47637/sienna.v7i1.2368

Pengembangan Arsitektur Hybrid LSTM dan Isolation Forest untuk Deteksi *Realtime* Serangan SQLi, *Brute-Force*, dan *Data Exfiltration* pada MariaDB

Khusnul Khotimah[a,1]*, Hartono[a,2], Rama Apriando[a,3], Faris[a,4]
[1*]khusnul.khotimah@umko.ac.id, [2]hartono@umko.ac.id, [3]rama.apriando@umko.ac.id, [4]faris@umko.ac.id
[a]Universitas Muhammadiyah Kotabumi, Lampung
*Korespondensi: email

Abstract

Cybersecurity threats targeting database systems continue to increase, particularly SQL injection (SQLi), brute-force attacks, and data exfiltration, which can compromise data confidentiality, integrity, and availability. Conventional rule-based detection approaches often struggle to identify evolving attack patterns and may generate high false-positive rates. This research aims to develop a real-time attack detection system for MariaDB databases using a hybrid architecture that combines Long Short-Term Memory (LSTM) and Isolation Forest to improve detection accuracy and adaptability in modern database environments. The study employed a Research and Development (R&D) approach consisting of three detection modules. The first module utilized a Bidirectional LSTM (Bi-LSTM) model to identify SQLi attacks through query sequence analysis. The second module implemented Isolation Forest and Rotated Isolation Forest algorithms to detect brute-force attacks based on access behavior patterns. The third module applied Isolation Forest to identify data exfiltration activities from traffic characteristics and data transfer behavior. Data preprocessing included feature engineering, tokenization, normalization, and model evaluation using confusion matrix, precision, recall, F1-score, and AUC metrics. Experimental results demonstrated that the Bi-LSTM model achieved 99.99% accuracy in SQLi detection. For brute-force detection, the standard Isolation Forest outperformed Rotated Isolation Forest with a recall of 99.94% and an F1-score of 99.61%. The data exfiltration detection module successfully identified 100% of simulated exfiltration attacks and achieved an overall accuracy of 94.92%. These findings indicate that the proposed hybrid LSTM–Isolation Forest architecture provides highly accurate and adaptive real-time attack detection, making it a viable solution for enhancing MariaDB database security against multiple attack vectors.

Status Artikel:
Diterima: 30-05-2026
Direvisi: 13-06-2026
Diterima: 14-06-2026

Kata Kunci:
MariaDB Security;
SQL Injection Detection;
Brute-Force Detection;
Data Exfiltration;
Hybrid LSTM-Isolation Forest.

PENDAHULUAN

MariaDB merupakan salah satu sistem manajemen basis data relasional yang banyak digunakan pada berbagai aplikasi modern, mulai dari layanan bisnis, pendidikan, hingga sistem

pemerintahan. Tingginya penggunaan MariaDB menjadikannya sebagai target utama berbagai serangan siber yang terus berkembang dari sisi teknik, skala, maupun kompleksitas. Berdasarkan berbagai penelitian sebelumnya (Bhardwaj dkk., 2022; Ghozali dkk., 2022; Hartono dkk., 2024; Stiawan dkk., 2023; Triloka dkk., 2022), terdapat tiga jenis serangan yang paling sering mengancam keamanan *database*, yaitu SQL Injection (SQLi), *brute-force*, dan *data exfiltration*. Ketiga serangan tersebut mampu mengeksploitasi celah keamanan aplikasi secara masif dan persisten sehingga berpotensi menyebabkan kebocoran data, gangguan layanan, hingga kerusakan sistem.

SQL Injection (SQLi) masih menjadi salah satu ancaman paling berbahaya terhadap *database* karena mampu memanipulasi *query* untuk memperoleh akses ilegal ke data sensitif. Menurut BitNinja Security (Molnar, 2021b), SQLi berkontribusi pada lebih dari 50% kasus serangan pada aplikasi web. Selain itu, serangan *brute-force* juga terus meningkat dan menjadi metode yang paling sering digunakan dalam berbagai kasus data *breach* (Faircloth dkk., 2022; Molnar, 2021a; Sharma dkk., 2024). Serangan ini dilakukan melalui percobaan *login* berulang secara otomatis untuk mendapatkan kredensial akses *database*. Kondisi yang lebih kritis terjadi pada *data exfiltration* karena serangan ini umumnya berjalan secara tersembunyi dan sulit dideteksi oleh sistem keamanan tradisional (Mundt & Baier, 2024; Security Magazine Staff, 2023). Pelaku dapat memindahkan data sensitif keluar sistem tanpa menimbulkan indikasi serangan yang jelas sehingga menyebabkan kerugian besar bagi organisasi.

Di sisi lain, pendekatan keamanan tradisional seperti Web Application Firewall (WAF) dan *signature-based detection* memiliki keterbatasan dalam mendeteksi pola serangan baru dan serangan yang terus berubah secara dinamis. Gartner melaporkan bahwa sistem berbasis pola statis gagal mendeteksi sekitar 56% serangan *zero-day* (F5, Inc., 2023; Goh dkk., 2023). Kondisi tersebut menunjukkan bahwa metode konvensional tidak lagi memadai untuk menghadapi ancaman keamanan *database* modern yang semakin adaptif dan kompleks.

Penelitian sebelumnya menunjukkan bahwa metode LSTM telah banyak digunakan untuk mendeteksi serangan SQL Injection (SQLi) karena mampu mengenali pola *sekuens query* dengan baik (Ghozali dkk., 2022; Liu & Dai, 2024; Stiawan dkk., 2023). Namun, pendekatan tersebut masih menghasilkan *false positive* yang cukup signifikan dalam proses deteksi. Di sisi lain, Isolation Forest digunakan pada beberapa penelitian untuk mendeteksi anomali *brute-force*, lonjakan lalu lintas jaringan, serta aktivitas *data exfiltration* mencurigakan (Marteau dkk., 2017; Mundt & Baier, 2024; Mykhaylova dkk., 2023; Vaccari dkk., 2021). Meskipun memiliki kemampuan yang baik dalam mendeteksi *outlier*, metode tersebut masih kurang optimal dalam menangkap pola temporal pada log aktivitas secara berurutan. Berdasarkan keterbatasan tersebut, penelitian ini mengusulkan pengembangan arsitektur hybrid yang mengintegrasikan LSTM dan Isolation Forest dalam satu sistem terpadu. Pendekatan hybrid ini diharapkan mampu meningkatkan akurasi deteksi, mengurangi *false positive* melalui mekanisme konfirmasi dua lapis, serta mendukung proses deteksi *real-time* dengan *latency* rendah pada lingkungan MariaDB.

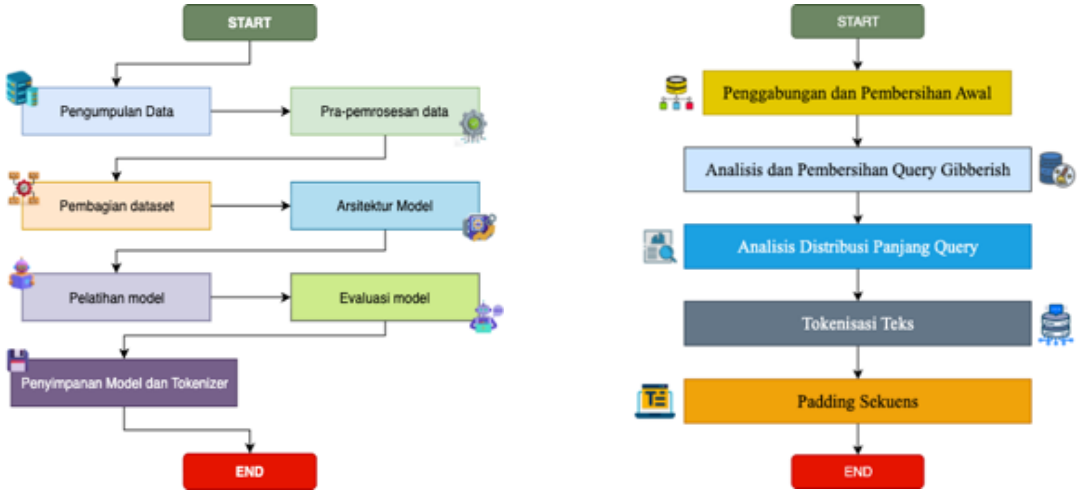
Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pengembangan arsitektur hybrid berbasis Long Short-Term Memory (LSTM) dan Isolation Forest untuk mendeteksi serangan secara *real-time* pada MariaDB. LSTM digunakan untuk menganalisis pola temporal *query* dan mendeteksi SQL Injection, sedangkan Isolation Forest dimanfaatkan untuk mengidentifikasi anomali perilaku akses yang berkaitan dengan *brute-force* dan *data exfiltration*.

Arsitektur hybrid ini dirancang agar mampu memberikan deteksi yang lebih adaptif, akurat, dan responsif terhadap ancaman keamanan siber modern, sehingga dapat meningkatkan keamanan *database* MariaDB secara signifikan.

METHODS

SQL Injection

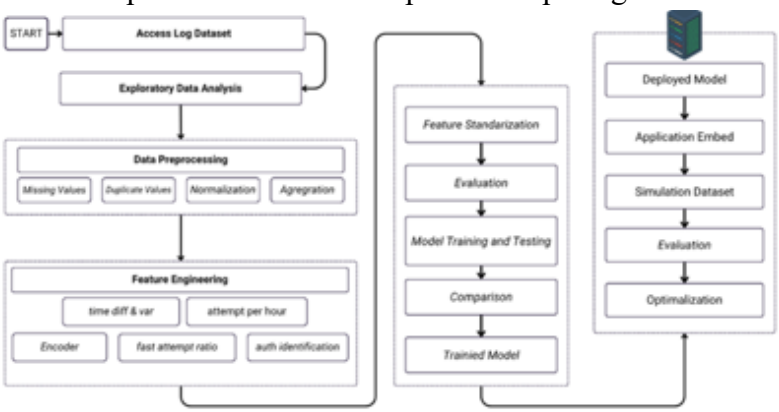
Penelitian ini menggunakan model Bi-LSTM berbasis DL untuk mendeteksi serangan SQLi di MariaDB. Selain itu, penelitian ini juga menerapkan NLP pada *preprocessing* untuk semakin meningkatkan tingkat akurasi. Tahapan penelitian dilakukan melalui beberapa proses yaitu pengumpulan data, pra-pemrosesan data, pembangunan arsitektur model, pelatihan model, serta evaluasi performa.



Gambar 1 Metode Penelitian dan tahapan pra-pemrosesan data Deteksi SQLi

Brute Force

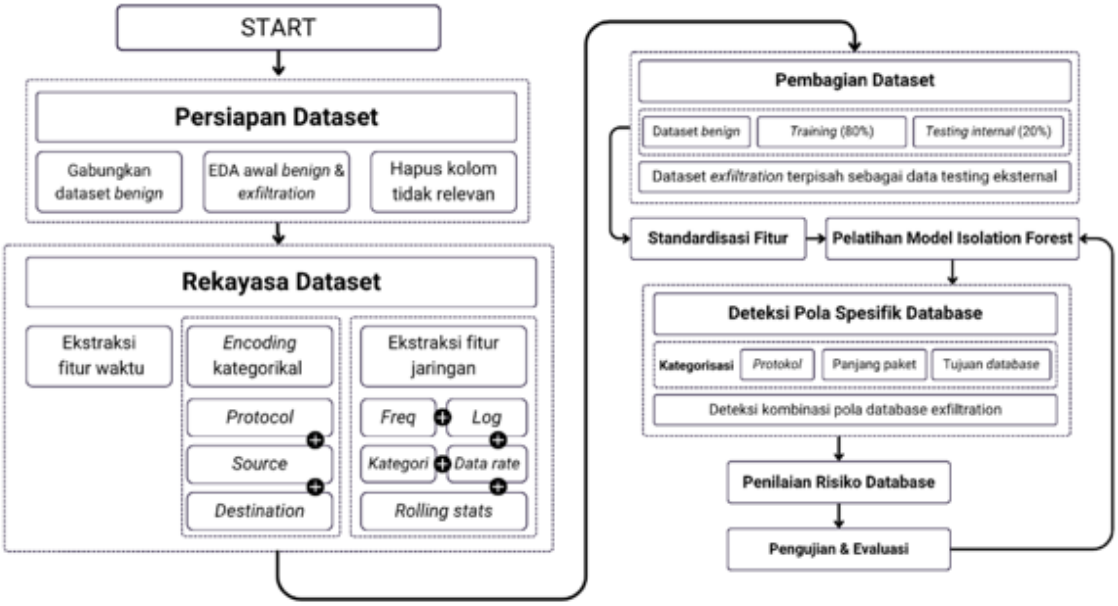
Penelitian ini memilih pendekatan *unsupervised learning* dengan algoritma IF untuk mendeteksi serangan *brute force* pada MariaDB. IF dipilih karena telah diterapkan secara luas untuk mendeteksi anomali pada berbagai domain keamanan siber, seperti deteksi intrusi jaringan, *malware*, anomali lalu lintas web, keamanan email, dan perlindungan infrastruktur kritis. Pemilihan metodologi ini dilandasi oleh kebutuhan untuk mendeteksi anomali secara efektif pada data log yang berukuran besar, kompleks, dan tidak selalu memiliki label eksplisit. Dengan pendekatan ini, penelitian diharapkan menghasilkan sistem deteksi yang *robust*, sistematis, serta dapat direproduksi. Alur implementasi metode dapat dilihat pada gambar 2.



Gambar 2 Alur Metode Penelitian Deteksi Serangan *Brute Force*

Data Exfiltration

Penelitian ini membangun dan mengevaluasi sistem deteksi anomali *data exfiltration* dengan fokus pada lalu lintas *database* MariaDB menggunakan algoritma IF. Berdasarkan klasifikasi umum, pendekatan yang digunakan dalam penelitian ini termasuk dalam kategori tindakan balasan detektif (*detective counter measures*), karena berfokus pada identifikasi upaya *exfiltration* saat sedang berlangsung, bukan pada pencegahan di tahap awal. Metodologi ini mencakup deskripsi *dataset* yang digunakan, tahapan pra-pemrosesan data, detail algoritma deteksi anomali, serta metrik evaluasi kinerja model. Gambar 11 menunjukkan bagaimana metode di penelitian ini diterapkan.



Gambar 3 Metode Penelitian Data Exfiltration

RESULTS AND DISCUSSION

Results

SQL Injection

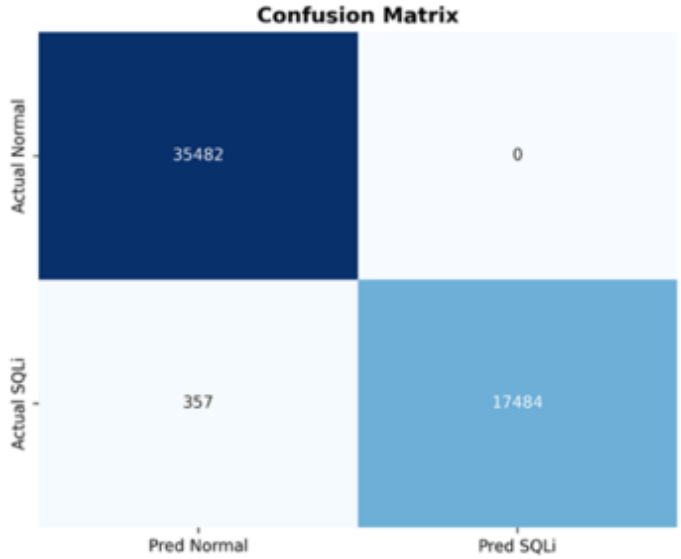
Model Bi-LSTM untuk deteksi SQL Injection dilatih menggunakan *dataset* terkurasi sebanyak 266.615 *query* dan dievaluasi pada 53.323 sampel data uji. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-score untuk mengukur kemampuan model dalam membedakan *query* normal dan *query* berbahaya secara akurat. Hasil pengujian menunjukkan bahwa model mampu mencapai performa yang sangat tinggi yakni ≥ 98% dengan tingkat kesalahan klasifikasi yang rendah. Rincian hasil evaluasi model Bi-LSTM pada data uji SQLi disajikan pada Tabel 1.

Tabel 1 Hasil Evaluasi Model Bi-LSTM pada Data Uji SQLi

Metrik	Nilai	Interprestasi
Overall Accuracy	99.99%	Persentase total *query* (Normal + SQLi) yang diklasifikasikan dengan benar oleh model.
Precision (Normal, 0)	99%	Dari semua *query* yang diprediksi sebagai Normal, 99% benar-benar *query* Normal. Tingkat *False Positive* sangat rendah.

Metrik	Nilai	Interprestasi
Recall (Normal, 0)	100%	Dari seluruh *query* Normal yang sebenarnya, 100% berhasil dideteksi. Model tidak melewatkan *query* Normal (*False Negative* rendah).
F1-Score (Normal, 0)	99%	Rata-rata harmonik antara Precision dan Recall untuk kelas Normal, menunjukkan keseimbangan performa.
Precision (SQLi, 1)	99%	Dari semua *query* yang diprediksi sebagai SQL Injection, 99% benar-benar serangan SQLi.
Recall (SQLi, 1)	98%	Dari seluruh *query* SQLi yang sebenarnya, 98% berhasil dideteksi. *False Negative* sangat rendah, menandakan keamanan terjaga.
F1-Score (SQLi, 1)	99%	Rata-rata harmonik antara *Precision* dan *Recall* untuk kelas SQLi, menunjukkan performa seimbang.
Macro Avg	99%	Rata-rata metrik (*Precision*, *Recall*, F1) di semua kelas, memberikan gambaran keseluruhan performa.
Weighted Avg	99%	Rata-rata metrik berbobot sesuai jumlah *query* tiap kelas, mencerminkan performa keseluruhan dengan memperhitungkan distribusi label.

Hasil evaluasi disajikan dalam bentuk *classification report*, *confusion matrix*, dan visualisasi persentase klasifikasi (Gambar 4). *Confusion matrix* menunjukkan nilai *true positive*, *true negative*, *false positive*, dan *false negative* dalam proses klasifikasi *query*. Visualisasi persentase klasifikasi memperlihatkan dominasi prediksi benar pada diagonal utama dan kesalahan klasifikasi yang sangat rendah, sehingga menunjukkan bahwa model mampu membedakan *query* Normal dan SQL Injection dengan sangat baik.



Gambar 4 *Confusion Matrix SQL Injection*

Brute Force

Model Isolation Forest dilatih menggunakan hasil *data cleaning* sebanyak 6.958.042 baris unik untuk mendeteksi aktivitas *brute-force* pada MariaDB dan menghasilkan nilai akurasi 99% (IF) dan 98% (RIF). Evaluasi dilakukan dengan 100.000 *data dummy* untuk membandingkan performa Standard Isolation Forest (IF) dan Rotated Isolation Forest (RIF) dalam skenario *real-*

time. Hasil pengujian menunjukkan bahwa kedua model memiliki performa yang sangat tinggi, namun Standard IF memberikan hasil yang lebih stabil dan akurat dibandingkan RIF. Rincian hasil evaluasi Standard IF disajikan pada Tabel 2 berikut.

Tabel 2 Metrik Evaluasi dan *Confusion Matrix Standard IF*

	precision	***recall***	***f1-score***	***support***
Benign	0.991493	0.899571	0.943298	3.663980e+05
Brute Force	0.992869	0.999448	0.996148	5.126052e+06
accuracy	0.992786	0.992786	0.992786	9.927856e-01
macro avg	0.992181	0.949510	0.969723	5.492450e+06
weighted avg	0.992777	0.992786	0.992622	5.492450e+06

Metric	***Value***
Accuracy	0.992786
AUC	0.949510
F1 Score	0.996148
Precision	0.992869
Recall (Sensitivity)	0.999448
Specificity	0.899571

RIF memiliki performa yang tinggi juga tetapi jumlah *false negative* pada RIF lebih besar dibandingkan *Standard IF*, sehingga tingkat sensitivitasnya sedikit menurun. Rincian hasil evaluasi *Rotated IF* disajikan pada Tabel 3.

Tabel 3 Metrik Evaluasi dan *Confusion Matrix Rotated IF*

	precision	***recall***	***f1-score***	***support***
Benign	0.920117	0.900223	0.910062	3.663980e+05
Brute Force	0.992879	0.994414	0.993646	5.126052e+06
accuracy	0.988130	0.988130	0.988130	9.881303e-01
macro avg	0.956498	0.947318	0.951854	5.492450e+06
weighted avg	0.988025	0.988130	0.988070	5.492450e+06

Metric	***Value***
Accuracy	**0.9881**
AUC	**0.9473**
F1 Score	**0.9936**
Precision	**0.9929**
Recall (Sensitivity)	**0.9944**
Specificity	**0.9002**

Data Exfiltration

Model Isolation Forest untuk deteksi anomali *data exfiltration* dilatih menggunakan 17.500 *data train* dan 4.000 *data test*. Hasil pengujian pada *test set internal* (*benign*) menunjukkan akurasi sebesar 94,92%, dari seluruh prediksi yang dikategorikan sebagai anomali, hanya sebagian kecil yang benar-benar merupakan kasus *data exfiltration* dalam *dataset* ini. Selain itu, pengujian pada *dataset exfiltration* menunjukkan rata-rata skor risiko sebesar 72,96, dengan 54 dari 67 rekaman memiliki skor risiko ≥ 70. Hasil ini menunjukkan bahwa model mampu mengidentifikasi pola transfer data mencurigakan dengan tingkat sensitivitas yang tinggi. Rincian hasil klasifikasi model Isolation Forest disajikan pada Gambar 5.

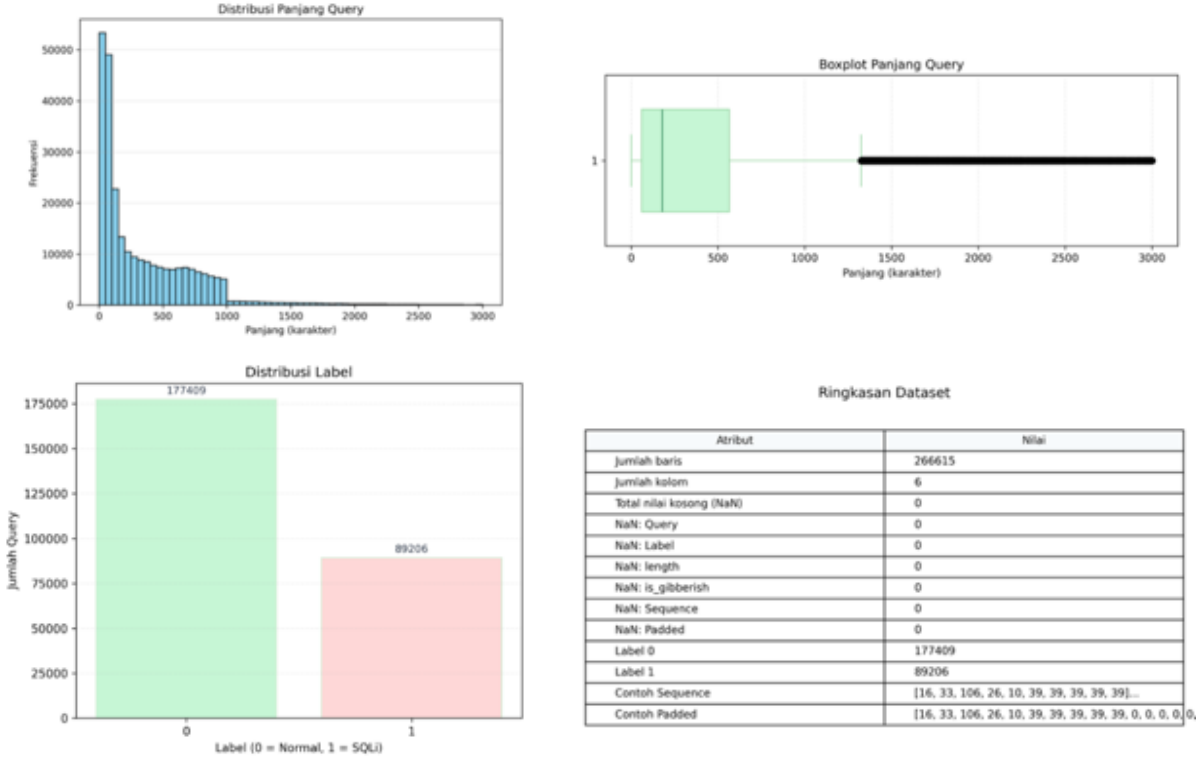
	precision	recall	f1-score	support
Benign	1.0000	0.9484	0.9735	4417
Exfiltration	0.2271	1.0000	0.3702	67
accuracy			0.9492	4484
macro avg	0.6136	0.9742	0.6718	4484
weighted avg	0.9885	0.9492	0.9645	4484

Gambar 5 *Classification Report* Hasil Prediksi Model IF

Discussion

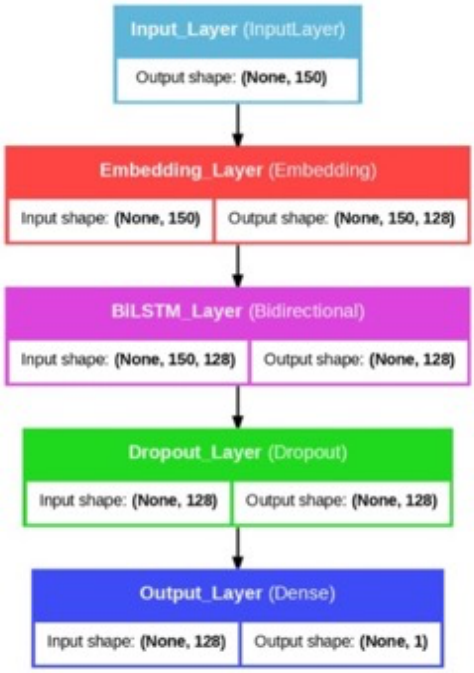
SQL Injection

Sebelum proses pelatihan dilakukan, *dataset* mentah terlebih dahulu melalui tahapan pra-pemrosesan untuk memastikan kualitas dan konsistensi data. Tahapan ini mencakup penggabungan beberapa *dataset*, pembersihan *query* yang tidak relevan seperti *gibberish* pada *query* normal, tokenisasi, serta *padding sequence*. Proses tersebut bertujuan agar model menerima data yang lebih terstruktur, bersih, dan optimal selama pelatihan maupun evaluasi. Hasil visualisasi *dataset* setelah tahapan pra-pemrosesan disajikan pada gambar berikut.



Gambar 6 Visualisasi Dataset setelah Melakukan Tahapan Pra-Pemrosesan Data

Model Bi-LSTM dibangun dengan arsitektur sebagai berikut: *input layer* dengan dimensi (150), *embedding layer* dengan *output* dimensi 128, lapisan Bi-LSTM dengan 64 unit, *dropout layer* dengan rate 0.5, serta *dense output layer* dengan aktivasi *sigmoid* untuk klasifikasi biner. Model dikompilasi menggunakan *loss function* binary_crossentropy, *optimizer* Adam, dan *metrik accuracy*. Proses pelatihan berlangsung selama 5 *epoch* dengan *batch size* 256 dan validation split 0.1.



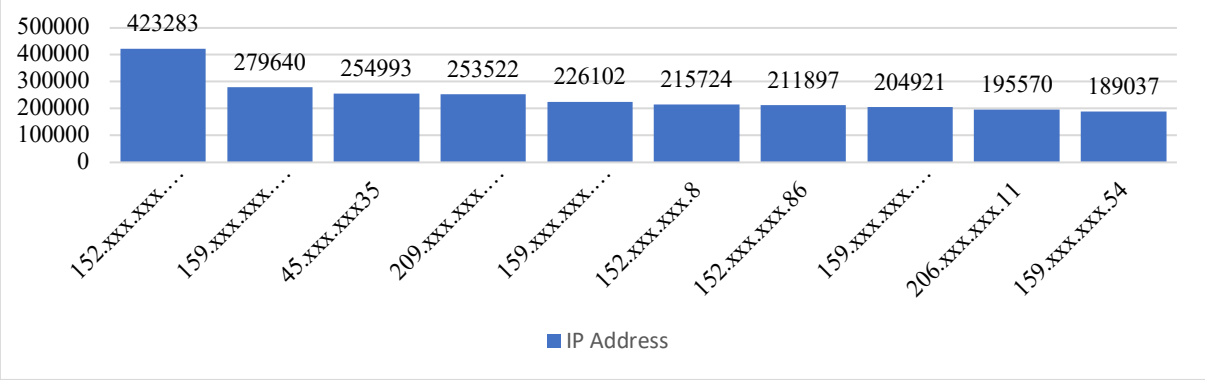
Gambar 7 Penerapan pada Arsitektur Jaringan Bi-LSTM untuk Deteksi Serangan SQL Injection

Hasil evaluasi model Bi-LSTM yang ditampilkan pada Tabel 1 dan divisualisasikan dalam Gambar 4 memberikan gambaran kinerja model secara komprehensif terhadap data uji yang terdiri dari 53.323 sampel. Secara umum, model menunjukkan kemampuan yang sangat baik dalam membedakan *query* Normal dan SQL Injection (SQLi). Akurasi keseluruhan model tercatat sebesar 0,99, menandakan bahwa 99% *query* pada data uji berhasil diklasifikasikan dengan benar. Nilai *macro average* dan *weighted average* untuk *precision*, *recall*, dan f1-score yang konsisten di angka 0,99 menunjukkan stabilitas performa model pada kedua kelas, tanpa bias terhadap salah satu kelas. Lebih rinci, pada kelas 0 (Normal), *precision* sebesar 0,99 mengindikasikan bahwa dari seluruh *query* yang diprediksi sebagai Normal, 99% memang benar-benar Normal. *Recall* sebesar 1,00 memperlihatkan bahwa seluruh *query* Normal dalam data uji berhasil dikenali oleh model tanpa ada kasus FN.

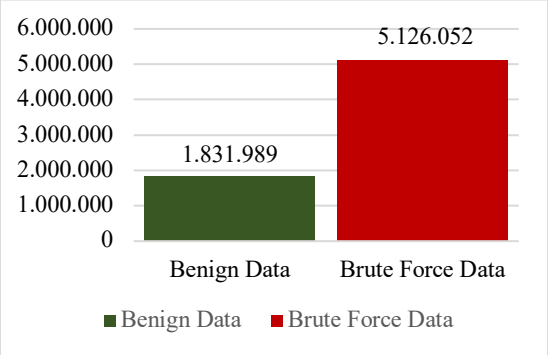
Kemampuan model tersebut menunjukkan bahwa model Bi-LSTM mampu mempelajari pola *sekuens query* SQL secara efektif sehingga dapat membedakan aktivitas normal dan serangan SQL Injection dengan tingkat kesalahan yang sangat rendah. Hasil ini sejalan dengan penelitian Ghozali dkk. (2022), Liu dan Dai (2024), serta Stiawan dkk. (2023) yang menyatakan bahwa arsitektur LSTM memiliki kemampuan yang baik dalam menangkap dependensi temporal pada *query* SQL. Namun demikian, pada penelitian ini performa yang diperoleh mencapai akurasi 99% dengan keseimbangan *precision* dan *recall* yang tinggi, menunjukkan bahwa pendekatan yang diterapkan mampu menghasilkan deteksi yang stabil dan minim *false positive*.

Brute Force

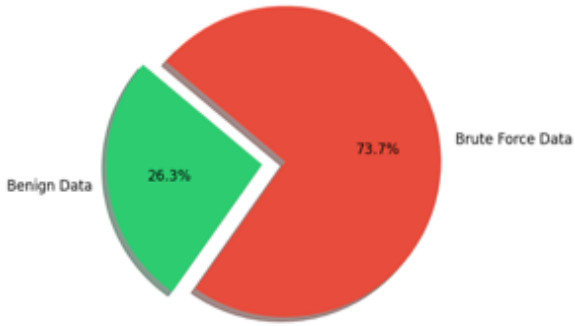
Terdapat total 25 domain pada *access log*. Dari total 25 domain tersebut, terdapat 1 domain yang diakses hingga 4.604.678 kali. Dengan membandingkan hasil terhadap ambang batas (*threshold*), kode mengidentifikasi pasangan IP–path yang menunjukkan anomali intensitas atau frekuensi *login* tinggi sebagai indikasi *brute force attack*, menghasilkan tiga keluaran: data normal, data serangan, dan statistik agregat per *IP–path*. Proses ini mendapatkan 1.831.989 data normal dan 5.126.052 serangan seperti yang ditunjukkan pada gambar berikut. Distribusi 10 alamat IP terbanyak akses dapat dilihat pada gambar 5 berikut ini.



Gambar 8 Distribusi 10 Alamat IP Terbanyak Akses



Gambar 9 Distribusi Total Insiden: *Benign* vs *Brute force*



Gambar 10 Proporsi *Data Benign* dan *Brute-force* dalam *Dataset*

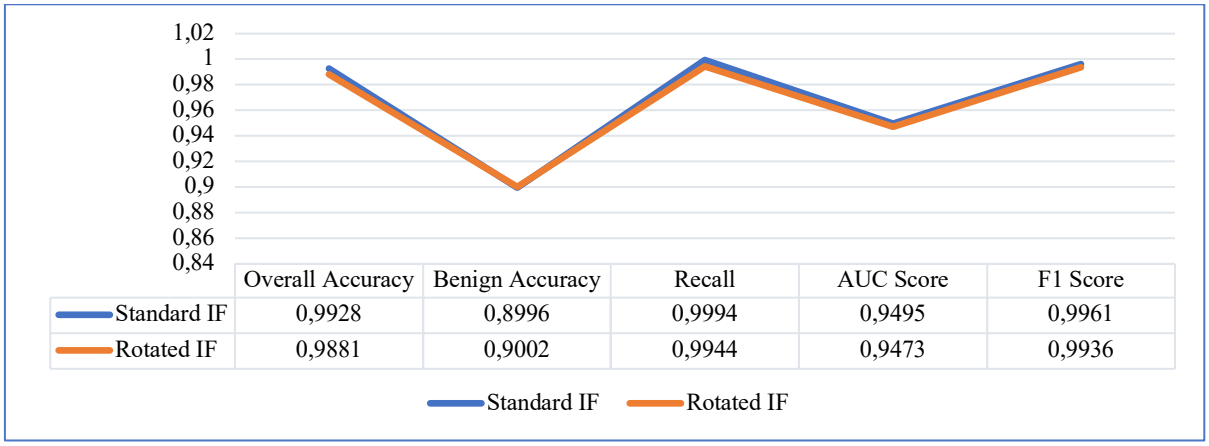
Hasil penghitungan fitur statistik temporal yang digunakan untuk menganalisis pola serangan *brute force* berdasarkan log aktivitas. Fitur-fitur seperti rata-rata waktu antar percobaan, variansi waktu, dan rasio percobaan cepat menjadi indikator penting dalam membedakan perilaku serangan dengan aktivitas normal.

Tabel 4 Hasil Perhitungan Fitur Statistik Temporal pada Analisis Serangan *Brute Force*

Statistic	Total Attempts	Avg Time Diff	Min Time Diff	Fast Attempts	Time Variance	Time Span Hours	Attempts Per Hour	Fast Attempt Rasio
count	85,65	25,111	25,111	85,650	16,468	85,650	85,650	85,650
mean	61.94	165,029	107,405	58.89	4.15×10^{10}	36.85	36.25	0.0249
std	2,293	278,950	279,564	2,284.	1.61×10^{11}	99.99	116.75	0.1150
min	1.0	0.0	0.0	0.0	0.0	0.0	0.0014	0.0
25%	1.0	10,063	25.0	0.0	4.97×10^{8}	0.0	15.69	0.0
50%	1.0	82,725	5,720.0	0.0	9.46×10^{9}	0.0	40.00	0.0
75%	2.0	196,927	51,317.0	0.0	2.52×10^{10}	0.33	40.00	0.0

Statistic	Total Attempts	Avg Time Diff	Min Time Diff	Fast Attempts	Time Variance	Time Span Hours	Attempts Per Hour	Fast Attempt Rasio
max	210,26	5,151,442	5,151,442	210,251	5.62×10^{12}	1,430.96	8,775.17	0.999992

Model IF dibangun dengan parameter n_estimators=175 untuk jumlah pohon, contamination=0.1 sebagai proporsi data anomali, max_samples='auto' agar setiap pohon menggunakan hingga 256 sampel, max_features=1.0 untuk memakai seluruh fitur, bootstrap=False tanpa pengembalian sampel, random_state=6 guna menjaga konsistensi hasil, verbose=0 agar proses berjalan tanpa log, dan n_jobs=-1 untuk memanfaatkan seluruh CPU. Model kemudian dilatih pada data ter-standardisasi menggunakan metode fit. Selanjutnya, Eksperimen dilakukan dengan melatih model Standard IF dan Rotated IF pada *data training benign* yang telah distandardisasi. Kedua model dievaluasi performanya dalam mengklasifikasikan data uji, yang terdiri dari campuran *data benign* dan *brute force*. Tabel 4 merangkum metrik evaluasi kunci untuk kedua model. Hasil evaluasi menunjukkan bahwa kedua pendekatan tersebut mampu mendeteksi serangan dengan tingkat akurasi yang sangat baik. Pada beberapa metrik tertentu, *Rotated* IF sedikit lebih unggul dalam mengidentifikasi *data benign*, sedangkan Standard IF menunjukkan keunggulan dalam *overall accuracy*. Temuan ini mengindikasikan pemilihan model dapat disesuaikan dengan prioritas sistem keamanan, apakah fokus pada deteksi atau meminimalkan kesalahan klasifikasi pada data normal.



Gambar 11 Perbandingan Kinerja Model Standard Isolation Forest dan Rotated Isolation Forest

Hasil evaluasi menunjukkan bahwa model IF mampu membedakan dengan baik antara aktivitas normal dan serangan *brute force*. *Precision* yang tinggi (0,9929) menunjukkan bahwa sebagian besar deteksi serangan oleh model memang benar-benar serangan, sehingga risiko alarm palsu (*false positive*) relatif rendah. *Recall* atau sensitivitas yang sangat tinggi (0,9944) menegaskan bahwa model hampir selalu mampu menangkap serangan *brute force* yang terjadi, sehingga kemungkinan adanya serangan yang terlewat (*false negative*) sangat kecil. Spesifisitas sebesar 0,9002 menandakan model juga cukup andal dalam mengenali aktivitas normal, meskipun masih ada sejumlah aktivitas normal yang salah dikategorikan sebagai serangan. Dengan demikian, model IF tidak hanya efektif mendeteksi serangan, tetapi juga dapat diandalkan dalam meminimalkan alarm palsu yang membebani tim keamanan.

Tabel 5 Hasil Evaluasi Kinerja Model Deteksi Serangan *Brute Force*

Metrik	Nilai	Interpretasi
True Positive (TP)	5.123.224	Jumlah serangan *brute force* yang berhasil terdeteksi dengan benar oleh model. Angka tinggi menunjukkan model sangat efektif mengenali aktivitas berbahaya.
False Negative (FN)	2.828	Kasus serangan *brute force* yang tidak terdeteksi (terlewat) oleh model. Nilai kecil menandakan tingkat kelulusan serangan sangat rendah.
True Negative (TN)	329.601	Jumlah aktivitas normal (*benign*) yang terklasifikasi benar sebagai aman. Ini menunjukkan kemampuan model dalam mengenali lalu lintas normal.
False Positive (FP)	36.797	Aktivitas normal yang salah diklasifikasikan sebagai serangan. Nilai ini perlu dikendalikan agar tidak menimbulkan peringatan palsu (*false alarm*).
Recall (*Brute Force Detection Rate*)	99.94% (0.9994)	Persentase serangan *brute force* yang berhasil terdeteksi dari total serangan sebenarnya. Nilai hampir sempurna menunjukkan sensitivitas model yang sangat tinggi.
Precision (*Brute Force*)	99.29% (0.9929)	Persentase prediksi serangan yang benar-benar serangan. Nilai tinggi menandakan model jarang memberikan peringatan palsu.
Specificity (*True Negative Rate*)	89.96% (0.8996)	Kemampuan model dalam mengenali aktivitas normal dengan benar. Masih terdapat sebagian kecil aktivitas normal yang terdeteksi salah sebagai serangan.
Benign Classification Accuracy	89.96% (0.8996)	Akurasi pengenalan aktivitas normal (setara *specificity*). Metrik ini penting untuk menilai keseimbangan performa antara deteksi serangan dan pengenalan *trafik* normal.
F1-Score	0.9961	Rata-rata harmonik antara *Precision* dan *Recall*. Nilai sangat tinggi ini menunjukkan keseimbangan optimal antara deteksi yang akurat dan minim kesalahan.
AUC (*Area Under Curve*)	0.949510	Menunjukkan kemampuan model membedakan antara serangan dan *trafik* normal. Nilai mendekati 1 menunjukkan performa klasifikasi yang sangat baik.
Accuracy (Keseluruhan)	0.992786 (99.28%)	Persentase keseluruhan prediksi yang benar (baik serangan maupun normal). Nilai hampir sempurna menunjukkan model sangat kuat secara umum.

Secara keseluruhan, metrik-metrik ini menunjukkan bahwa model *standard IF* yang dikembangkan efektif dalam mendeteksi serangan *brute force* pada *dataset* log akses web ini. *Detection Rate* yang tinggi (99.94%) adalah kekuatan utama, yang sangat penting dalam aplikasi keamanan siber untuk meminimalkan serangan yang tidak terdeteksi. Tingkat *False Positive* relatif terkendali mengingat volume data dan kompleksitas pola *trafik*.

Temuan penting dari pengujian *brute force* adalah bahwa Standard Isolation Forest memberikan performa yang lebih baik dibandingkan Rotated Isolation Forest, terutama pada metrik *recall* dan akurasi keseluruhan. Hasil ini menunjukkan bahwa pendekatan Isolation Forest standar lebih sesuai untuk karakteristik data log akses yang digunakan dalam penelitian ini.

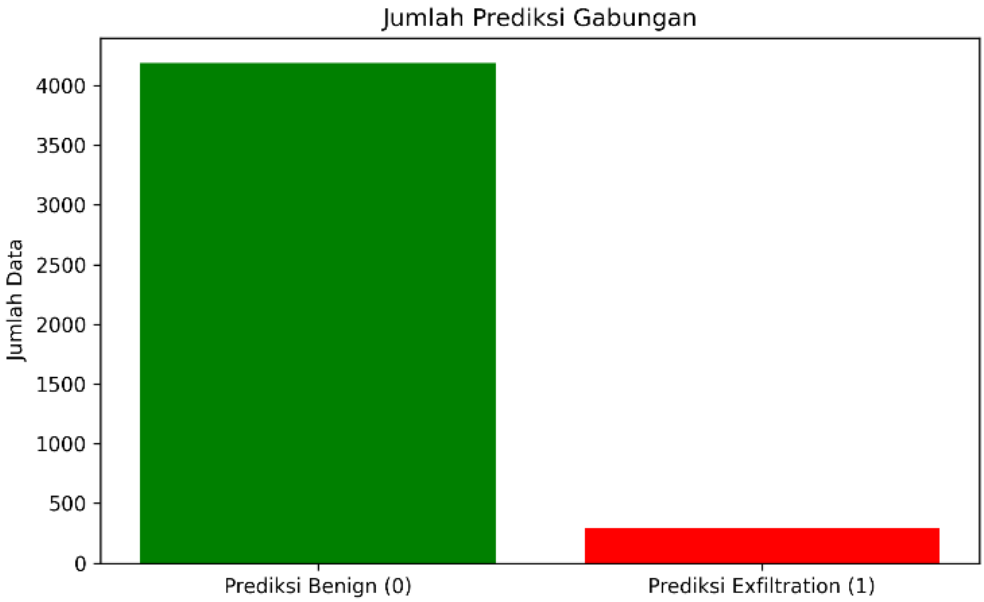
Temuan tersebut mendukung penelitian sebelumnya oleh Marteau dkk. (2017), Vaccari dkk. (2021), dan Mykhaylova dkk. (2023) yang menunjukkan bahwa Isolation Forest efektif dalam mendeteksi aktivitas anomali dan serangan berbasis perilaku. Tingginya nilai *recall* sebesar 99,94% juga menunjukkan bahwa model mampu meminimalkan kemungkinan serangan *brute force* yang tidak terdeteksi, yang merupakan aspek penting dalam sistem keamanan siber.

Data Exfiltration

Data pelatihan memiliki jumlah yang lebih besar dibandingkan data pengujian, yaitu sekitar 17.500 data untuk *train set* dan 4.000 data untuk *test set*. Model ini dikonfigurasi dengan *contamination*=0.05, menunjukkan asumsi bahwa sekitar 5% data dalam *training set* mungkin merupakan *outlier*. Hasil prediksi pada *test set* internal (*benign*) menunjukkan akurasi sebesar 0.9484 (94.84%). *Confusion Matrix* untuk pengujian. Hasil evaluasi menunjukkan bahwa dari 4.417 *record benign*, sebanyak 4.189 *record* berhasil diklasifikasikan dengan benar sebagai non-anomali (*True Negative*), sedangkan 228 *record* salah diklasifikasikan sebagai anomali (*False Positive*). Nilai *False Positive* Rate (FPR) pada test set internal tercatat sebesar 5,16% (228/4417). Hasil prediksi model Isolation Forest (IF) menunjukkan bahwa seluruh 67 *record exfiltration* berhasil terdeteksi sebagai anomali, menghasilkan *Detection Rate* (*True Positive Rate*) sebesar 1.0000 (100%).

Secara keseluruhan, evaluasi gabungan terhadap *test* set internal dan eksternal menghasilkan akurasi total sebesar 0.9492 (94,92%), menandakan kinerja model yang cukup andal dalam mendeteksi anomali jaringan. Berdasarkan *Classification Report* gabungan, kinerja per kelas dapat dirinci sebagai berikut:

a. Kelas *Benign* (0): *Precision* 1.00, *Recall* 0.95, F1-score 0.97
b. Kelas *Exfiltration* (1): *Precision* 0.23, *Recall* 1.00, F1-score 0.37



Gambar 12 Jumlah Prediksi Gabungan pada Kelas Benign dan Exfiltration

Gambar 12 memperlihatkan hasil prediksi model terhadap dua kelas utama, yaitu *Benign (0)* dan *Exfiltration (1)*. Mayoritas data diprediksi sebagai *Benign* dengan jumlah lebih dari 4000 sampel, sedangkan prediksi untuk kelas *Exfiltration* jauh lebih sedikit, hanya sekitar 300 sampel. Distribusi yang tidak seimbang ini menunjukkan bahwa aktivitas normal (*benign traffic*) masih

mendominasi *dataset*, sementara aktivitas anomali berupa *data exfiltration* terjadi relatif jarang. Hal ini menggambarkan tantangan umum dalam deteksi anomali jaringan, yaitu ketidakseimbangan kelas yang signifikan.

Hasil pengujian secara eksplisit menunjukkan bahwa model menghasilkan 4189 *True Negatives* (TN) atau *data Benign* yang berhasil diprediksi sebagai normal, 228 *False Positives* (FP) yaitu data Benign yang keliru diklasifikasikan sebagai anomali, 0 *False Negatives* (FN) atau tidak ada *data Exfiltration* yang salah diklasifikasikan sebagai normal, serta 67 *True Positives* (TP) yakni *data Exfiltration* yang berhasil terdeteksi sebagai anomali. Temuan ini menegaskan bahwa model Isolation Forest (IF) memiliki kemampuan yang sangat baik dalam mendeteksi anomali dengan *Detection Rate* mencapai 100%, namun masih menghadapi tantangan dalam membedakan antara anomali sebenarnya dengan *outlier* pada data normal. Hal ini tercermin dari nilai *False Positive* Rate (FPR) sebesar 5,16% dan *Precision* untuk kelas *Exfiltration* yang hanya 0,23. Rendahnya nilai *Precision* disebabkan oleh jumlah *False Positive* yang jauh lebih tinggi dibandingkan *True Positive* (228 FP berbanding 67 TP).

Pengujian pada *dataset exfiltration* (*df_test_ex*) menghasilkan *dataframe* hasil yang mencakup prediksi IF, kategori protokol, pola panjang paket, kategori tujuan, pola spesifik *database*, tingkat risiko *database*, dan skor risiko *database*. Semua 67 *record* dalam *dataset exfiltration* terdeteksi sebagai anomali oleh IF.

Tabel 6 Distribusi Protokol *Database*

No	*Protocol Category*	*Prediction*			*Database Risk Score*	
		Count	*Sum*	*Mean*	*Mean*	*Max*
1	DATABASE_DIRECT_CRITICAL_RISK	2	2	1.0	100.0	100
2	DATABASE_ENCRYPTED_VERY_HIGH_RISK	2	2	1.0	99.0	99
3	DATABASE_EXPORT_HTTP_HIGH_RISK	4	4	1.0	87.0	89
4	DATABASE_TCP_HIGH_RISK	1	1	1.0	89.0	89
5	DATABASE_WEB_HIGH_RISK	1	1	1.0	87.0	87

Analisis terhadap distribusi kategori protokol *database* menunjukkan bahwa rekaman *exfiltration* tersebar pada beberapa kelompok risiko. Berdasarkan Tabel 6, kategori dengan tingkat risiko tertinggi adalah *DATABASE_DIRECT_CRITICAL_RISK*, yang umumnya terkait dengan penggunaan protokol seperti MYSQL/MARIADB atau koneksi langsung menuju *port* standar 3306. Pada kategori ini, skor prediksi anomali rata-rata tercatat sebesar 1.0, sementara skor risiko *database* mencapai nilai maksimum, yaitu 100. Temuan ini sejalan dengan karakteristik lalu lintas *exfiltration* yang bersifat langsung, jelas, dan memiliki tingkat urgensi yang tinggi. Selain itu, analisis terhadap distribusi pola *database* mengungkapkan sejumlah pola yang muncul dalam *data exfiltration* simulasi, antara lain *LARGE_DATA_TRANSFER*, *DIRECT_DB_CONNECTION, CLOUD_STORAGE_TARGET*, *ENCRYPTED_BULK_TRANSFER*, dan *SUSPICIOUS_DESTINATION*. Kombinasi pola yang melibatkan koneksi langsung ke *database* disertai dengan transfer data berukuran besar, terutama apabila diarahkan ke tujuan yang dikategorikan mencurigakan, menunjukkan skor risiko tertinggi, yaitu mencapai 100.

Temuan utama pada skenario *data exfiltration* adalah kemampuan model Isolation Forest mendeteksi seluruh *data exfiltration* yang diuji dengan *Detection Rate* sebesar 100%. Hasil ini

menunjukkan bahwa karakteristik anomali pada aktivitas *exfiltration* berhasil dikenali dengan baik oleh model melalui pola transfer data, tujuan koneksi, dan perilaku lalu lintas jaringan. Temuan tersebut sejalan dengan penelitian Mundt dan Baier (2024) yang menyatakan bahwa pendekatan berbasis deteksi anomali efektif digunakan untuk mengidentifikasi aktivitas *data exfiltration* yang sulit dideteksi oleh metode berbasis aturan. Meskipun demikian, nilai *precision* yang masih rendah menunjukkan adanya *false positive* pada sebagian *trafik* normal sehingga optimasi lebih lanjut tetap diperlukan untuk meningkatkan ketepatan klasifikasi anomali.

SIMPULAN

Modul pertama menangani deteksi SQL Injection melalui Bi-LSTM. *Dataset* multi-sumber diproses melalui pembersihan duplikasi, penanganan *missing* values, normalisasi *query*, tokenisasi, dan *padding*. Arsitektur final terdiri atas *embedding layer*, Bi-LSTM *layer*, *dropout*, dan *dense layer*. Evaluasi menunjukkan performa sangat tinggi dengan akurasi 99,99%, *precision* 99,99%, *recall* 99,99%, dan *F1-score* 99,99%, menegaskan kemampuan model dalam membedakan *query benign* dan *malicious* secara nyaris sempurna. Model dan tokenizer disimpan untuk kebutuhan *inference real-time*.

Modul kedua memfokuskan pada deteksi *brute-force* menggunakan Isolation Forest standar dan *Rotated* Isolation Forest. *Dataset* log akses berskala besar dikurasi menjadi 1.831.989 entri *benign* dan 5.126.052 *entri brute-force* melalui *pipeline preprocessing* dan rekayasa fitur perilaku akses. IF standar memberikan performa terbaik dengan *Recall* 99,94%, *Precision* 99,29%, F1-*Score* 99,61%, AUC 0,9495, dan akurasi 99,28%. *Confusion matrix* menunjukkan TP = 5.123.224, TN = 329.601, FP = 36.797, dan FN = 2.828. Sebagai perbandingan, Rotated Isolation Forest menghasilkan *Recall* 99,44%, *Precision* 99,29%, *F1-Score* 99,36%, AUC 0,9473, dan akurasi 98,81%, dengan TP = 5.097.416, TN = 329.840, FP = 36.558, dan FN = 28.636. Hasil ini mengonfirmasi bahwa IF standar lebih stabil dan presisi pada konteks *brute-force* berskala besar.

Modul ketiga menerapkan Isolation Forest untuk deteksi *data exfiltration*. Dataset benign digabungkan kemudian diperkaya melalui EDA dan rekayasa fitur, mencakup profil waktu, frekuensi, panjang paket, protokol, dan tujuan trafik. Model menghasilkan akurasi internal 94,84% pada *benign data*, dan 100% *detection rate* pada *simulated exfiltration traffic* (67 kasus). Akurasi gabungan mencapai 94,92% dengan *confusion matrix*: *True Positive* = 67, *False Positive* = 228, *True Negative* = 4.189, dan *False Negative* = 0. Analisis lanjutan menunjukkan bahwa protokol langsung (MySQL/MariaDB), transfer data besar, dan tujuan eksternal seperti *cloud storage* memiliki *database risk score* tertinggi. Secara keseluruhan, penelitian ini membuktikan bahwa arsitektur hybrid LSTM–Isolation Forest mampu memberikan deteksi yang akurat, adaptif, dan dapat dioperasionalkan secara real-time untuk tiga vektor serangan kritis terhadap MariaDB. Dengan tingkat akurasi sangat tinggi pada SQLi dan *brute-force*, serta tingkat deteksi sempurna untuk skenario eksfiltrasi data, arsitektur ini menawarkan landasan kuat bagi pengembangan sistem pemantauan keamanan siber generasi berikutnya.

REFERENSI

Bhardwaj, A., Chandok, S. S., Bagnawar, A., Mishra, S., & Uplaonkar, D. (2022). Detection of Cyber Attacks: XSS, SQLI, Phishing Attacks and Detecting Intrusion Using Machine Learning Algorithms. *2022 IEEE Global Conference on Computing, Power and*

Communication Technologies (GlobConPT), 1–6. https://doi.org/10.1109/GlobConPT57482.2022.9877770

F5, Inc. (2023). *Gartner\textsuperscript® Report: Market Guide for Cloud Web Application and API Protection*. F5, Inc. https://www.f5.com/c/apcj/2023/asset/gartner-report-market-guide-for-cloud-web-application-and-api-protection

Faircloth, C., Hartzell, G., Callahan, N., & Bhunia, S. (2022). A Study on Brute Force Attack on T-Mobile Leading to SIM-Hijacking and Identity-Theft. *2022 IEEE World AI IoT Congress (AIIoT)*, 501–507. https://doi.org/10.1109/AIIoT54504.2022.9817175

Ghozali, I., Asy'ari, M. F., Triarjo, S., Ramadhani, H. M., Studiawan, H., & Shiddiqi, A. M. (2022). A Novel SQL Injection Detection Using Bi-LSTM and TF-IDF. *2022 7th International Conference on Information and Network Technologies (ICINT)*, 16–22. https://doi.org/10.1109/ICINT54892.2022.9939254

Goh, J. R., Wang, S. S., Harel, Y., & Toh, G. (2023). Predictive Taxonomy Analytics (LASSO): Predicting Outcome Types of Cyber Breach. *Journal of Cybersecurity*, *9*(1), tyad015. https://doi.org/10.1093/cybsec/tyad015

Hartono, H., Wijaya, R. A., & Khotimah, K. (2024). Development of Detection and Mitigation of Advanced Persistent Threats Using Artificial Intelligence and Multi-Layer Security on Cloud Computing Infrastructure. *International Journal of Artificial Intelligence Research*, *8*(2), 194–211. https://doi.org/10.29099/ijair.v8i2.1026

Liu, Y., & Dai, Y. (2024). Deep Learning in Cybersecurity: A Hybrid BERT–LSTM Network for SQL Injection Attack Detection. *IET Information Security*, *2024*(1), 5565950. https://doi.org/10.1049/ise2.5565950

Marteau, P.-F., Soheily-Khah, S., & Béchet, N. (2017). *Hybrid Isolation Forest – Application to Intrusion Detection*. http://arxiv.org/abs/1705.03800

Molnar, A. (2021a). *The Most Common Types of Cyberattacks #3 – Brute Force Attacks*. BitNinja Security. https://bitninja.com/blog/the-most-common-types-of-cyberattacks-3-brute-force-attacks/

Molnar, A. (2021b). *The Most Common Types of Cyberattacks #4 – SQL Injection*. BitNinja Security. https://bitninja.com/blog/the-most-common-types-of-cyberattacks-4-sql-injection-attacks/

Mundt, M., & Baier, H. (2024). Enhancing Incident Management by an Improved Understanding of Data Exfiltration: Definition, Evaluation, Review. Dalam S. Goel & P. R. Nunes de Souza (Ed.), *Digital Forensics and Cyber Crime* (hlm. 33–57). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-xxxx-x_3

Mykhaylova, O., Shtypka, A., & Fedynyshyn, T. (2023). An Isolation Forest-Based Approach for Brute Force Attack Detection. *Cybersecurity and Information Technologies*, *23*(4), 45–56. https://doi.org/10.2478/cait-2023-0045

Security Magazine Staff. (2023). *There was a 39% surge in data exfiltration cyberattacks in 2023*. Security Magazine. https://www.securitymagazine.com/articles/100359-there-was-a-39-surge-in-data-exfiltration-cyberattacks-in-2023

Sharma, P., Tanwar, S., & Kukreja, V. (2024). Implementation of Brute Force Attack. Dalam *Emerging Trends in IoT and Computing Technologies*. CRC Press.

Stiawan, D., Bardadi, A., Afifah, N., Melinda, L., Heryanto, A., Septian, T. W., & others. (2023). An Improved LSTM-PCA Ensemble Classifier for SQL Injection and XSS Attack Detection. *Computer Systems Science and Engineering*, *46*(2), 1759–1774. https://doi.org/10.32604/csse.2023.036061

Triloka, J., Hartono, H., & Sutedi, S. (2022). Detection of SQL Injection Attack Using Machine Learning Based On Natural Language Processing. *International Journal of Artificial Intelligence Research*, *6*(2). https://doi.org/10.29099/ijair.v6i2.355

Vaccari, I., Narteni, S., Aiello, M., Mongelli, M., & Cambiaso, E. (2021). Exploiting Internet of Things Protocols for Malicious Data Exfiltration Activities. *IEEE Access*, *9*, 104261–104280. https://doi.org/10.1109/ACCESS.2021.3099576