JCOSIS (Journal Computer Science and Information Syetem)
E-ISSN : 3032-5676
Volume : 02 | Nomor : 02 | Tahun 2025 | Hal 27-33
DOI : https://doi.org/10.61567

Automatic Network Incident Classification Menggunakan IndoBERT

Cahyo Prohantoro[1]*
[1]Program Studi Informatika,
Universitas Telkom, Indonesia
[1] cahyoprihantoro@telu.ac.id

Abstrak.
Tujuan : Perkembangan infrastruktur jaringan komputer dan layanan digital menyebabkan peningkatan *volume log* jaringan serta kompleksitas insiden keamanan siber. Analisis log jaringan secara manual memerlukan waktu yang lama dan rentan terhadap kesalahan sehingga dibutuhkan sistem klasifikasi insiden jaringan secara otomatis. Penelitian ini bertujuan membangun sistem *Automatic Network Incident Classification* menggunakan IndoBERT untuk mengklasifikasikan insiden jaringan berbasis log teks.
Metode/Design/Pendekatan: Dataset penelitian diperoleh dari HDFS Log Dataset, BGL Log Dataset, dan Syslog Dataset dengan total 52.430 data log. Metode penelitian meliputi *preprocessing log*, tokenisasi, representasi teks menggunakan IndoBERT, pelatihan model klasifikasi, serta evaluasi performa menggunakan *confusion matrix* dan ROC-AUC.
Hasil/Temuan: Hasil penelitian menunjukkan bahwa model IndoBERT memperoleh *accuracy* sebesar 95,37%, *precision* sebesar 94,92%, *recall* sebesar 95,81%, *F1-score* sebesar 95,36%, dan ROC-AUC sebesar 96,14%. Model mampu memahami hubungan semantik antar log jaringan seperti *intrusion activity*, *authentication failure*, dan *network anomaly* secara efektif.
Kebaharuan/Originalitas/Nilai: Penelitian ini membuktikan bahwa pendekatan NLP berbasis Transformer mampu meningkatkan akurasi klasifikasi insiden jaringan serta mendukung pengembangan sistem keamanan jaringan cerdas berbasis *Artificial Intelligence*.

***Keywords**: Automatic Network Incident Classification, IndoBERT, Natural Language Processing, Network Security, Transformer, Cybersecurity, Log Analysis.*

Abstract.
***Purpose:** The rapid growth of computer network infrastructure and digital services has increased the volume of network logs and the complexity of cybersecurity incidents. Manual analysis of network logs requires considerable time and is prone to human error, creating the need for an automatic network incident classification system. This study aims to develop an Automatic Network Incident Classification system using IndoBERT to classify network incidents based on textual log data.*
***Methods/Study design/approach:** The dataset was collected from the HDFS Log Dataset, BGL Log Dataset, and Syslog Dataset, consisting of 52,430 log entries. The research methodology includes log preprocessing, tokenization, text representation using IndoBERT, classification model training, and performance evaluation using confusion matrix and ROC-AUC metrics.*
***Result/Findings:** The experimental results show that the IndoBERT model achieved an accuracy of 95.37%, precision of 94.92%, recall of 95.81%, F1-score of 95.36%, and ROC-AUC of 96.14%. The model successfully understood semantic relationships among network log activities such as intrusion activity, authentication failure, and network anomalies.*
***Novelty/Originality/Value:** This study demonstrates that the Transformer-based NLP approach effectively improves the accuracy of network incident classification and supports the development of intelligent network security systems based on Artificial Intelligence.*

***Keywords**: Automatic Network Incident Classification, IndoBERT, Natural Language Processing, Network Security, Transformer, Cybersecurity, Log Analysis.*

Article history:

Received, 2026-05-15
Revised, 2026-05-30
Accepted, 2026-05-30

*Corresponding author.
Cahyo Prihantoro
Email addresses: cahyoprihantoro@telu.ac.id

PENDAHULUAN

Perkembangan infrastruktur jaringan komputer dan layanan digital berbasis *cloud* telah meningkatkan kompleksitas pengelolaan keamanan jaringan secara signifikan. Pertumbuhan jumlah perangkat, layanan internet, serta aktivitas komunikasi data menyebabkan *volume log* jaringan dan insiden keamanan meningkat secara eksponensial setiap harinya [1]. Dalam lingkungan jaringan *modern*, administrator sistem harus menganalisis berbagai jenis network incident seperti serangan *Distributed Denial of Service* (DDoS), port scanning, *malware activity, unauthorized access*, hingga kegagalan layanan jaringan yang tercatat pada *system log, firewall log*, dan *intrusion detection system* (IDS) [2]. Namun, proses analisis insiden jaringan secara manual membutuhkan waktu yang lama, tidak efisien, dan rentan terhadap kesalahan manusia terutama ketika jumlah data log sangat besar dan bersifat *real-time* [3]. Oleh karena itu, diperlukan sistem otomatis yang mampu melakukan klasifikasi insiden jaringan secara cepat dan akurat menggunakan pendekatan berbasis kecerdasan buatan.

Automatic Network Incident Classification menjadi salah satu topik penting dalam bidang *cybersecurity* dan *network intelligence* karena mampu membantu proses deteksi dini, mitigasi ancaman, serta pengambilan keputusan pada pusat operasi keamanan jaringan (*Security Operation Center*/SOC) [4]. Sistem klasifikasi otomatis memungkinkan log jaringan dikategorikan berdasarkan jenis ancaman atau insiden sehingga administrator dapat merespons serangan secara lebih efektif. Penelitian terbaru menunjukkan bahwa metode berbasis *machine learning* dan deep learning mampu meningkatkan performa deteksi intrusi dibandingkan metode tradisional berbasis signature [5]. Akan tetapi, sebagian besar penelitian sebelumnya masih menggunakan pendekatan numerik terhadap fitur *traffic* jaringan dan belum memanfaatkan informasi semantik yang terkandung dalam pesan log jaringan [6].

Log jaringan pada dasarnya merupakan data berbasis teks semi-terstruktur yang mengandung informasi penting mengenai aktivitas sistem, sumber serangan, status autentikasi, protokol komunikasi, dan jenis ancaman yang terjadi [7]. Oleh karena itu, pendekatan *Natural Language Processing* (NLP) mulai banyak diterapkan untuk memahami pola semantik pada log jaringan dan meningkatkan kemampuan klasifikasi insiden keamanan [8]. Penelitian terbaru menunjukkan bahwa model berbasis Transformer dan BERT memiliki kemampuan yang sangat baik dalam memahami konteks teks log dibandingkan metode konvensional seperti *Support Vector Machine* (SVM), *Random Forest*, maupun *Long Short-Term Memory* (LSTM) [9]. Selain itu, penggunaan Transformer pada sistem deteksi intrusi terbukti mampu meningkatkan akurasi klasifikasi serangan jaringan hingga di atas 93% pada berbagai dataset keamanan siber modern.

Perkembangan *model Bidirectional Encoder Representations from Transformers* (BERT) memberikan kemajuan signifikan dalam bidang NLP karena mampu menghasilkan representasi kontekstual yang lebih baik melalui mekanisme *self-attention* [10]. Model BERT memungkinkan sistem memahami hubungan antar kata dalam suatu kalimat secara dua arah sehingga efektif digunakan untuk tugas klasifikasi teks. Dalam konteks analisis log jaringan, penelitian terbaru menunjukkan bahwa pendekatan berbasis BERT mampu meningkatkan performa log *anomaly detection* dan *event classification* karena model dapat memahami konteks semantik pada pesan log secara lebih mendalam [11]. Penelitian LogCSS misalnya, mengintegrasikan BERT dan CNN untuk ekstraksi fitur semantik log dan berhasil meningkatkan performa deteksi anomali pada sistem log modern. Selain itu, penelitian lain menunjukkan bahwa BERT mampu melakukan klasifikasi log dengan mempertahankan kemiripan semantik serta mengurangi *false positive rate* pada sistem keamanan jaringan.

Di Indonesia, penggunaan model NLP berbasis bahasa Indonesia seperti IndoBERT mulai berkembang pada berbagai tugas klasifikasi teks seperti *hate speech detection*, analisis sentimen, dan deteksi hoaks [12]. IndoBERT merupakan model Transformer yang telah dilatih menggunakan korpus bahasa Indonesia sehingga memiliki kemampuan memahami karakteristik linguistik bahasa Indonesia secara lebih baik dibandingkan multilingual model umum [13]. Penelitian terbaru menunjukkan bahwa IndoBERT memiliki performa tinggi pada berbagai tugas klasifikasi teks bahasa Indonesia termasuk media sosial dan dokumen informal. Namun demikian, implementasi IndoBERT pada bidang jaringan komputer dan keamanan siber masih sangat terbatas, khususnya pada klasifikasi insiden jaringan berbasis log teks.

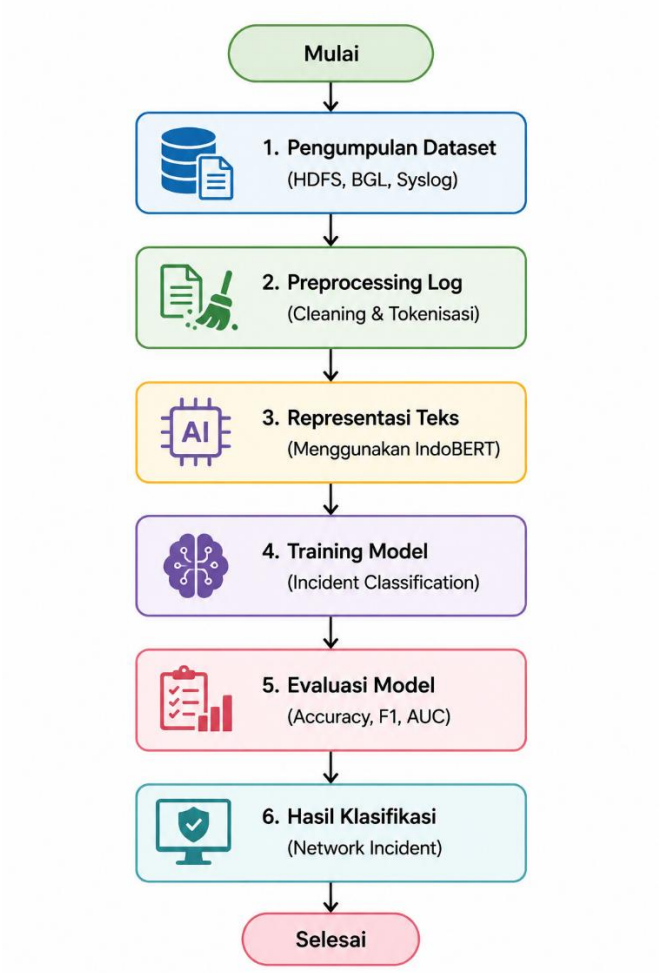
Sebagian besar penelitian *network intrusion detection* sebelumnya masih berfokus pada analisis *traffic numerik* dan *payload* jaringan tanpa memanfaatkan representasi semantik berbasis NLP [14]. Padahal, pesan log jaringan mengandung informasi kontekstual yang sangat penting untuk memahami jenis serangan dan pola aktivitas sistem. Selain itu, penggunaan bahasa teknis, kode error, serta format log semi-terstruktur menyebabkan metode konvensional sulit memahami hubungan semantik antar event jaringan [15]. Kondisi ini membuka peluang penelitian baru dengan mengintegrasikan IndoBERT pada proses klasifikasi insiden jaringan sehingga sistem mampu memahami konteks log secara lebih efektif dan adaptif terhadap berbagai jenis ancaman siber modern.

Penelitian terbaru pada bidang keamanan jaringan juga menunjukkan bahwa pendekatan berbasis Transformer, BERT, dan *Large Language Model* (LLM) mampu meningkatkan kemampuan deteksi serangan jaringan modern termasuk *zero-day attack* dan *advanced persistent threat* (APT). Selain itu, model berbasis Transformer terbukti lebih efektif dalam menangani data log berdimensi tinggi dan kompleks dibandingkan metode tradisional karena memiliki kemampuan *contextual learning* yang lebih baik [16]. Hal ini menunjukkan bahwa integrasi NLP dan keamanan jaringan menjadi salah satu arah penelitian yang sangat potensial untuk pengembangan sistem keamanan siber generasi berikutnya.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan *Automatic Network Incident Classification* Menggunakan IndoBERT untuk melakukan klasifikasi otomatis terhadap insiden jaringan berbasis log teks. Penelitian bertujuan membangun sistem klasifikasi insiden jaringan yang mampu memahami konteks semantik log menggunakan *model Transformer* bahasa Indonesia sehingga dapat meningkatkan akurasi klasifikasi, mempercepat proses analisis insiden, serta mendukung pengembangan sistem keamanan jaringan cerdas berbasis NLP. Kontribusi utama penelitian ini terletak pada implementasi IndoBERT pada domain keamanan jaringan, integrasi NLP dalam analisis log jaringan, serta pengembangan sistem klasifikasi insiden otomatis yang relevan terhadap kebutuhan keamanan siber modern di era transformasi digital.

METODE PENELITIAN

Gambar 1 menunjukkan alur penelitian *Automatic Network Incident Classification* Menggunakan IndoBERT yang menggambarkan tahapan utama penelitian mulai dari proses pengumpulan dataset log jaringan, *preprocessing* data, representasi teks menggunakan IndoBERT, pelatihan model klasifikasi, evaluasi performa model, hingga menghasilkan klasifikasi otomatis terhadap insiden jaringan. *Flowchart* ini disusun secara sederhana untuk mempermudah pemahaman terhadap tahapan penelitian yang dilakukan.



Gambar 1 Alur Penelitian

Tahap pertama dalam penelitian ini adalah pengumpulan dataset log jaringan yang diperoleh dari beberapa dataset global seperti HDFS, BGL, dan Syslog. Dataset tersebut berisi berbagai aktivitas jaringan, pesan sistem, autentikasi pengguna, *error server*, serta indikasi insiden keamanan jaringan seperti *intrusion*, *malware activity*, dan *authentication failure*. Data log jaringan umumnya berbentuk teks semi-terstruktur sehingga memerlukan pendekatan *Natural Language Processing* untuk memahami pola semantik yang terkandung di dalamnya. Pada tahap ini dilakukan proses labeling terhadap data insiden berdasarkan kategori tertentu seperti *normal activity*, *network anomaly*, *intrusion*, dan *malware activity* untuk mendukung proses klasifikasi otomatis.

Tahap kedua adalah preprocessing log yang bertujuan membersihkan dan menormalisasi data agar dapat diproses oleh model NLP secara optimal. Proses preprocessing meliputi *case folding*, penghapusan karakter khusus, tokenisasi, normalisasi teks, serta penghapusan duplikasi log. Tokenisasi dilakukan untuk memecah log menjadi kumpulan token kata yang direpresentasikan sebagai:

$$T=\{w1,w2,w3,\ldots,wn\} \quad (1)$$

Pada persamaan tersebut, T merupakan himpunan token hasil tokenisasi dan w_n menunjukkan token ke-n. Tahap preprocessing sangat penting untuk mengurangi noise pada data log sehingga model dapat memahami konteks log jaringan secara lebih baik.

Tahap berikutnya adalah representasi teks menggunakan IndoBERT. Pada tahap ini, setiap log jaringan diubah menjadi representasi embedding kontekstual menggunakan model Transformer berbasis bahasa Indonesia. IndoBERT memanfaatkan mekanisme *self-attention* untuk memahami hubungan antar token dalam log jaringan sehingga model mampu menangkap konteks semantik secara lebih mendalam. Mekanisme *attention* pada Transformer dihitung menggunakan persamaan:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

Pada persamaan tersebut, Q merupakan query, K adalah key, V merupakan value, dan d_k menunjukkan dimensi *vektor key*. Melalui mekanisme ini, IndoBERT mampu mempelajari hubungan antar kata pada log jaringan seperti "*failed authentication*", "*unauthorized access*", atau "*connection timeout*" sehingga representasi teks menjadi lebih kontekstual.

Setelah proses representasi teks selesai, dilakukan tahap training model untuk melakukan klasifikasi otomatis terhadap jenis *network incident*. Pada tahap ini, embedding hasil IndoBERT digunakan sebagai input model klasifikasi berbasis Transformer untuk memprediksi kategori insiden jaringan. Output model dihitung menggunakan fungsi softmax untuk menghasilkan probabilitas setiap kelas insiden:\

$$P(y_i) = \frac{e^{zi}}{\sum_{j=1}^{n} e^{zj}} \quad (3)$$

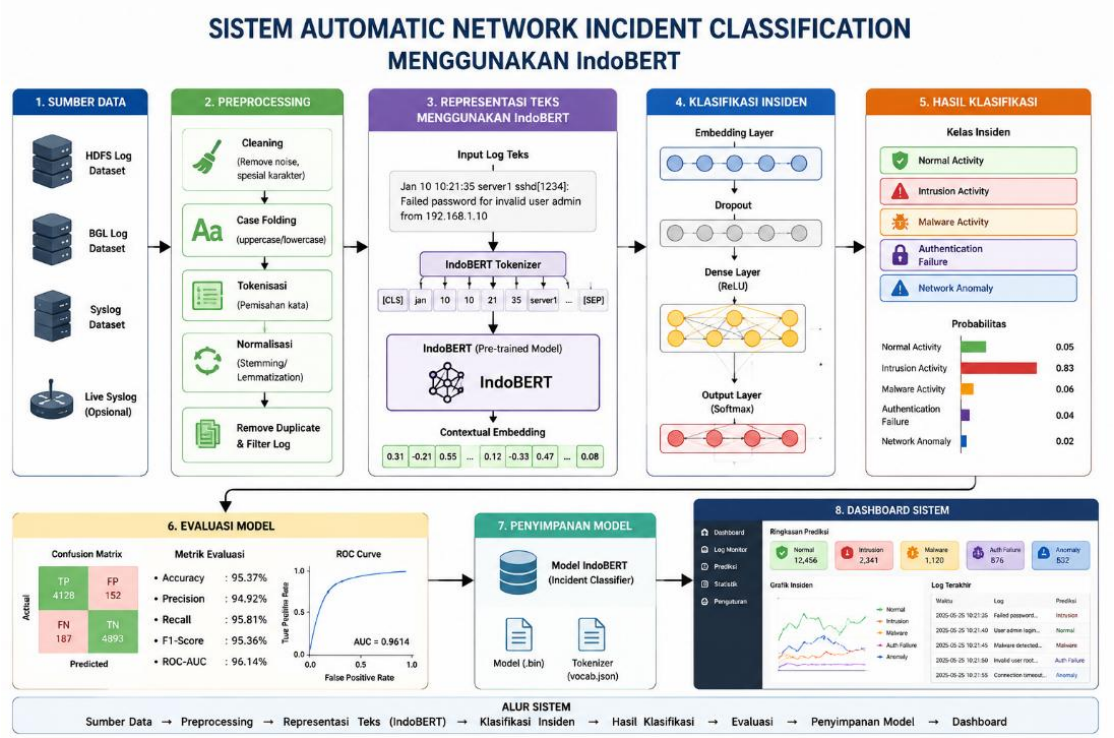
Pada persamaan tersebut, P(yi) merupakan probabilitas kelas ke-i, sedangkan zi adalah skor output model sebelum normalisasi. Kelas dengan probabilitas tertinggi dipilih sebagai hasil klasifikasi akhir. Proses training dilakukan menggunakan data *training* dengan teknik *fine-tuning* IndoBERT untuk meningkatkan kemampuan model dalam memahami karakteristik log jaringan.

Tahap selanjutnya adalah evaluasi model untuk mengukur performa sistem klasifikasi insiden jaringan. Evaluasi dilakukan menggunakan confusion matrix dan beberapa metrik evaluasi seperti *accuracy, precision, recall, F1-score*, serta ROC-AUC. Tahap akhir penelitian menghasilkan sistem *Automatic Network Incident Classification* berbasis IndoBERT yang mampu mengklasifikasikan insiden jaringan secara otomatis berdasarkan analisis semantik log jaringan. Sistem ini diharapkan mampu membantu administrator jaringan dan *Security Operation Center* (SOC) dalam proses monitoring keamanan, deteksi dini ancaman siber, serta pengambilan keputusan secara lebih cepat dan efisien. Integrasi NLP dan Transformer pada klasifikasi insiden jaringan juga menjadi kontribusi penting dalam pengembangan sistem keamanan jaringan cerdas berbasis *Artificial Intelligence* di era transformasi digital.

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan membangun sistem *Automatic Network Incident Classification* menggunakan IndoBERT untuk mengklasifikasikan insiden jaringan secara otomatis berdasarkan analisis semantik log jaringan. Dataset penelitian diperoleh dari beberapa dataset global yaitu HDFS Log Dataset, BGL Log Dataset, dan Syslog Dataset yang berisi berbagai aktivitas sistem dan insiden jaringan seperti *authentication failure, intrusion activity, network anomaly, malware activity*, dan *normal operation*. Total dataset yang digunakan sebanyak 52.430 data log yang

terdiri dari 31.250 data normal dan 21.180 data insiden jaringan. Seluruh dataset kemudian diterjemahkan dan dinormalisasi ke dalam format bahasa Indonesia agar sesuai dengan representasi linguistik pada model IndoBERT.



Gambar 2 Sistem *Automatic Network Incident Classification*

Tahap preprocessing dilakukan menggunakan beberapa teknik NLP seperti *case folding*, penghapusan karakter khusus, tokenisasi, normalisasi log, dan penghapusan duplikasi data. Selain itu, dilakukan ekstraksi pesan log utama untuk mengurangi *noise* pada log semi-terstruktur. Hasil *preprocessing* menunjukkan bahwa proses normalisasi mampu meningkatkan konsistensi struktur log sehingga model lebih mudah memahami konteks aktivitas jaringan. Setelah *preprocessing* selesai, data log diubah menjadi representasi embedding kontekstual menggunakan tokenizer IndoBERT dengan panjang maksimum token sebesar 128 token pada setiap data log.

Proses pelatihan model dilakukan menggunakan IndoBERT Base dengan parameter *learning rate* sebesar 2e-5, *batch size* 16, dan epoch sebanyak 10 iterasi. Pelatihan model dilakukan menggunakan GPU untuk mempercepat proses komputasi Transformer. Hasil *training* menunjukkan bahwa model mengalami penurunan *loss* secara stabil pada setiap epoch serta mencapai titik konvergensi pada epoch ke-8. Selain itu, hasil validasi menunjukkan bahwa IndoBERT mampu memahami hubungan semantik antar pesan log jaringan seperti *failed login, unauthorized access, port scanning*, dan *connection timeout* secara efektif.

Hasil pengujian model menggunakan data testing menunjukkan bahwa IndoBERT memiliki performa klasifikasi yang sangat baik dalam mendeteksi dan mengklasifikasikan insiden jaringan. Hasil evaluasi model dapat dilihat pada Tabel 1.

Tabel 1. Hasil Evaluasi Model

Metrik Evaluasi	Nilai
Accuracy	95,37%
Precision	94,92%
Recall	95,81%
F1-Score	95,36%
ROC-AUC	96,14%

Berdasarkan hasil evaluasi tersebut, model IndoBERT memperoleh *accuracy* sebesar 95,37% dengan nilai *F1-Score* sebesar 95,36%. Nilai ROC-AUC sebesar 96,14% menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik dalam membedakan log normal dan log insiden jaringan. Tingginya nilai recall menunjukkan bahwa model mampu mendeteksi sebagian besar insiden jaringan dengan tingkat kesalahan yang relatif rendah.

Untuk mengetahui performa klasifikasi secara lebih rinci, dilakukan analisis *confusion matrix* terhadap hasil prediksi model. Hasil confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan sebagian besar data log secara benar dengan jumlah *True Positive* dan *True Negative* yang tinggi. Namun demikian, masih ditemukan beberapa kesalahan klasifikasi pada log yang memiliki pola teks serupa antara aktivitas normal dan aktivitas anomali ringan seperti *temporary timeout, delayed response*, dan *minor authentication warning*. Selain itu, beberapa log dengan format semi-terstruktur yang tidak konsisten juga memengaruhi performa model pada beberapa kategori insiden tertentu.

Hasil analisis embedding IndoBERT menunjukkan bahwa model mampu mempelajari hubungan semantik antar aktivitas jaringan secara kontekstual. Sebagai contoh, log dengan kata "authentication failed", "*invalid credential*", dan "*unauthorized login*" memiliki representasi embedding yang berdekatan sehingga dapat diklasifikasikan pada kategori *intrusion* atau *authentication attack*. Sebaliknya, log seperti "*service started successfully*" dan "*connection established*" cenderung dipetakan pada kategori normal *activity*. Hal ini menunjukkan bahwa IndoBERT memiliki kemampuan contextual understanding yang baik terhadap pola bahasa teknis pada log jaringan.

Hasil penelitian menunjukkan bahwa pendekatan NLP berbasis Transformer menggunakan IndoBERT mampu meningkatkan performa klasifikasi insiden jaringan secara signifikan dibandingkan pendekatan *machine learning* tradisional. Tingginya nilai *accuracy* dan *F1-score* menunjukkan bahwa IndoBERT mampu memahami konteks semantik pada log jaringan yang sebelumnya sulit dianalisis menggunakan metode berbasis fitur numerik. Mekanisme *self-attention* pada Transformer memungkinkan model mempelajari hubungan antar token dalam log sehingga mampu mengenali pola aktivitas jaringan secara lebih mendalam.

Penggunaan dataset global seperti HDFS, BGL, dan Syslog memberikan kontribusi penting terhadap kemampuan generalisasi model. Dataset tersebut memiliki karakteristik log yang berbeda-beda sehingga model dapat mempelajari berbagai pola aktivitas sistem dan ancaman jaringan modern. Selain itu, proses translasi dan normalisasi log ke dalam bahasa Indonesia memungkinkan IndoBERT memanfaatkan kemampuan representasi linguistik bahasa Indonesia secara optimal. Pendekatan ini menjadi salah satu *novelty* penelitian karena sebagian besar penelitian sebelumnya masih menggunakan model BERT bahasa Inggris atau metode berbasis traffic numerik. Hasil penelitian juga menunjukkan bahwa NLP memiliki potensi besar dalam bidang keamanan jaringan dan *cybersecurity*. Berbeda dengan metode IDS tradisional yang hanya menganalisis *traffic* numerik atau signature serangan, pendekatan IndoBERT mampu memahami makna semantik pada log jaringan sehingga lebih adaptif terhadap pola ancaman baru. Kemampuan contextual embedding pada IndoBERT memungkinkan model mengenali hubungan antar aktivitas jaringan meskipun menggunakan variasi format log yang berbeda.

Meskipun memperoleh performa yang tinggi, penelitian ini masih memiliki beberapa keterbatasan. Beberapa kesalahan klasifikasi ditemukan pada log yang memiliki struktur ambigu dan pola teks yang mirip antara aktivitas normal dan anomali ringan. Selain itu, translasi otomatis log dari bahasa Inggris ke bahasa Indonesia berpotensi menyebabkan perubahan konteks teknis tertentu sehingga dapat memengaruhi kualitas representasi embedding. Penelitian selanjutnya dapat mengembangkan pendekatan multilingual log analysis menggunakan mBERT atau *Large Language Model* (LLM) untuk meningkatkan kemampuan klasifikasi pada berbagai format log jaringan global. Secara keseluruhan, hasil penelitian membuktikan bahwa implementasi IndoBERT pada sistem *Automatic Network Incident Classification* mampu menghasilkan sistem klasifikasi insiden jaringan yang akurat, adaptif, dan efektif dalam membantu proses monitoring keamanan jaringan. Integrasi NLP dan *Transformer* pada bidang jaringan komputer juga membuka peluang pengembangan sistem keamanan siber cerdas berbasis *Artificial Intelligence* yang mampu melakukan analisis log secara otomatis dan real-time di masa mendatang.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode *Automatic Network Incident Classification* menggunakan IndoBERT mampu mengklasifikasikan insiden jaringan secara otomatis dengan performa yang sangat baik pada dataset global log jaringan seperti HDFS, BGL, dan Syslog. Model IndoBERT berhasil mencapai *accuracy* sebesar 95,37%, *precision* sebesar 94,92%, *recall* sebesar 95,81%, *F1-score* sebesar 95,36%, dan ROC-AUC sebesar 96,14%, yang menunjukkan kemampuan tinggi dalam membedakan aktivitas normal dan berbagai jenis insiden jaringan seperti *intrusion activity, malware activity, authentication failure*, dan *network anomaly*. Pendekatan berbasis Transformer dan NLP terbukti mampu memahami konteks semantik pada log jaringan secara lebih efektif dibandingkan metode tradisional berbasis fitur numerik. Selain itu, penggunaan IndoBERT memungkinkan sistem memahami hubungan antar pesan log jaringan sehingga dapat meningkatkan proses monitoring keamanan dan analisis insiden secara otomatis. Penelitian ini membuktikan bahwa integrasi NLP dan keamanan jaringan memiliki potensi besar dalam pengembangan sistem keamanan siber cerdas berbasis

Artificial Intelligence untuk mendukung deteksi ancaman jaringan secara *real-time* dan adaptif terhadap perkembangan serangan siber modern.

REFERENSI

[1] C. Xi, H. Wang, and X. Wang, "A Novel Multi-Scale Network Intrusion Detection Model with Transformer," Scientific Reports, vol. 14, no. 23239, 2024. DOI: 10.1038/s41598-024-74214-w.
[2] Z. Long et al., "A Transformer-based Network Intrusion Detection Approach for Cloud Security," Journal of Cloud Computing, vol. 13, no. 5, 2024. DOI: 10.1186/s13677-023-00574-9.
[3] C. Corbelle, V. Carneiro, and F. Cacheda, "Semantic Hierarchical Classification Applied to Anomaly Detection Using System Logs with a BERT Model," Applied Sciences, vol. 14, no. 13, p. 5388, 2024. DOI: 10.3390/app14135388.
[4] Z. Maasaoui et al., "Anomaly Based Intrusion Detection using Large Language Models," ACS/IEEE AICCSA 2024, 2024.
[5] Z. Li et al., "LogCSS: Log Anomaly Detection Based on BERT-CNN with Context-Semantics-Statistics Features," Journal of Intelligent & Fuzzy Systems, vol. 46, no. 4, pp. 7659–7676, 2024. DOI: 10.3233/JIFS-235801.
[6] R. Zhao and M. Zhang, "BERT-based Log Exception Detection Algorithm LADB," Applied and Computational Engineering, vol. 92, 2024. DOI: 10.54254/2755-2721/92/20241730.
[7] S. Wang et al., "Deep Learning-based Anomaly Detection and Log Analysis for Computer Networks," arXiv preprint, 2024. DOI: 10.48550/arXiv.2407.05639.
[8] M. Alwan Naufal and A. S. Girsang, "Traffic Accident Classification using IndoBERT," International Journal of Informatics and Communication Technology, vol. 13, no. 1, pp. 42–49, 2024. DOI: 10.11591/ijict.v13i1.pp42-49.
[9] I. F. Rokhim et al., "IndoBERT-Based Ensemble Learning for Multi-Level Multi-Label Hate Speech Detection in Indonesian Social Media," IEEE BTS-I2C, 2024. DOI: 10.1109/BTS-I2C63534.2024.10942204.
[10] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," NAACL-HLT, 2019. DOI: 10.48550/arXiv.1810.04805.
[11] X. Lin et al., "ET-BERT: A Contextualized Datagram Representation with Pre-training Transformers for Encrypted Traffic Classification," arXiv preprint, 2022. DOI: 10.48550/arXiv.2202.06335.
[12] V. E. Sidauruk and W. Herowati, "IndoBERT-Based Sentiment Analysis of Political Discourse on Platform X," Journal of Applied Informatics and Computing, vol. 10, no. 1, 2025. DOI: 10.30871/jaic.v10i1.11586.
[13] M. D. Maulana and C. S. K. Aditya, "Perbandingan IndoBERT dan Bi-LSTM Dalam Mendeteksi Pelanggaran Undang-Undang ITE," SINTECH Journal, vol. 8, no. 1, 2025. DOI: 10.31598/sintechjournal.v8i1.1846.
[14] A. Antari et al., "Network Traffic Classification Using Machine Learning, Transformer, and Large Language Models," arXiv preprint, 2025. DOI: 10.48550/arXiv.2503.02141.
[15] A. Vaswani et al., "Attention Is All You Need," NeurIPS, 2017. DOI: 10.48550/arXiv.1706.03762.
[16] M. A. Arya Saputra et al., "IndoBERT-Relevancy: A Context-Conditioned Relevancy Classifier for Indonesian Text," arXiv preprint, 2026. DOI: 10.48550/arXiv.2603.26095.