Prosiding Seminar Nasional Ilmu Teknik
Volume 2 Nomor 2 Desember 2025
e-ISSN: 3063-5713; p-ISSN: 3063-4709, Hal 691-703
DOI: https://doi.org/10.61132/prosemnasproit.v2i2.82
Tersedia: https://prosiding.aritekin.or.id/index.php/PROSEMNASPROIT/

Deteksi Serangan pada *Internet of Vehicles* dengan Algoritma XGBoost dan *Feature Selection Information Gain*

Kurnianto Basuki[1*], Kurniabudi[2], Eko Arip Winanto[3]
[1,2,3] Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dinamika Bangsa, Indonesia
Email: kurniantobasuki679@gmail.com[1], kurniabudi@unama.ac.id[2], ekoaripwinanto@unama.ac.id[3]*

Jl. Jend. Sudirman, The Hok, Kec. Jambi Sel., Kota Jambi, Jambi 36138, Indonesia.
**Penulis Korespondensi: kurniantobasuki679@gmail.com*

Abstract. *The rapid development of the Internet of Vehicles (IoV) has introduced new security challenges, particularly in protecting Controller Area Network (CAN Bus) communications from cyberattacks such as Denial of Service (DoS) and spoofing attacks. This study proposes the implementation of the Extreme Gradient Boosting (XGBoost) algorithm combined with Information Gain feature selection to improve intrusion detection performance in IoV environments. The CICIoV2024 dataset, which represents both benign and malicious traffic, is used as the primary data source. The research process includes data integration, preprocessing, feature selection, data splitting, and model training using a 5-fold cross-validation approach. Experimental results demonstrate that the proposed model achieves outstanding performance, with accuracy, precision, recall, and F1-score exceeding 99.99%, and an Area Under Curve (AUC) value approaching 1.00. Furthermore, Information Gain successfully identifies the most influential CAN payload features, enhancing model efficiency without sacrificing accuracy. These findings confirm that the combination of Information Gain and XGBoost is highly effective for developing a fast, accurate, and efficient intrusion detection system in IoV networks.*

Keywords: *Extreme Gradient Boosting; Information Gain; Internet of Vehicles; Intrusion Detection System; Network Security*

Abstrak. Perkembangan pesat *Internet of Vehicles (IoV)* menghadirkan tantangan serius dalam aspek keamanan, khususnya pada komunikasi *Controller Area Network (CAN Bus)* yang rentan terhadap serangan siber seperti *Denial of Service (DoS)* dan *spoofing*. Penelitian ini bertujuan untuk menerapkan algoritma *Extreme Gradient Boosting (XGBoost)* yang dikombinasikan dengan metode *feature selection Information Gain* untuk meningkatkan performa deteksi serangan pada lingkungan IoV. Dataset CICIoV2024 digunakan sebagai sumber data utama yang merepresentasikan *trafik* normal dan *trafik* serangan. Tahapan penelitian meliputi integrasi data, *preprocessing*, seleksi fitur, pembagian data, serta pelatihan model menggunakan metode *5-fold cross-validation*. Hasil eksperimen menunjukkan bahwa model mampu mencapai performa sangat tinggi dengan nilai *accuracy*, *precision*, *recall*, dan *F1-score* di atas 99,99%, serta nilai *Area Under Curve (AUC)* mendekati 1,00. Selain itu, metode *Information Gain* berhasil mengidentifikasi fitur *payload CAN* yang paling berpengaruh sehingga meningkatkan efisiensi model tanpa menurunkan akurasi. Temuan ini membuktikan bahwa kombinasi *Information Gain* dan *XGBoost* efektif dalam membangun sistem *Intrusion Detection System (IDS)* yang cepat, akurat, dan efisien untuk jaringan IoV.

Kata kunci: *Extreme Gradient Boosting*; *Information Gain*; *Internet of Vehicles*; *Intrusion Detection System*; Keamanan Jaringan

1. LATAR BELAKANG

Internet of Vehicles (IoV) merupakan pengembangan dari *Internet of Things* (IoT) yang memungkinkan kendaraan saling terhubung dan berkomunikasi dengan infrastruktur, kendaraan lain, serta sistem eksternal untuk meningkatkan keselamatan dan efisiensi transportasi (Alcaraz et al., 2018; Kaiwartya et al., 2016; Janbi, 2025). Meskipun memberikan berbagai manfaat seperti peningkatan keselamatan dan efisiensi transportasi, IoV juga menghadirkan risiko keamanan yang signifikan. Salah satu komponen utama dalam sistem

Naskah Masuk: 10 Oktober 2025; Revisi: 26 Oktober 2025; Diterima: 25 November 2025; Terbit: 30 Desember 2025;

kendaraan modern adalah *Controller Area Network* (CAN Bus), yang berfungsi sebagai media komunikasi antar *Electronic Control Unit* (ECU) dan menjadi tulang punggung pertukaran data pada kendaraan modern, namun dirancang tanpa mekanisme keamanan bawaan (Miller & Valasek, 2015; Woo et al., 2015; Janbi, 2025). Namun, CAN Bus tidak dirancang dengan mekanisme keamanan bawaan, sehingga rentan terhadap berbagai serangan siber.

Serangan seperti *Denial of Service* (DoS) dan *spoofing* dapat menyebabkan gangguan serius pada sistem kendaraan, mulai dari kesalahan pembacaan sensor hingga kegagalan fungsi kritis kendaraan, sehingga berpotensi membahayakan keselamatan pengguna dan stabilitas sistem kendaraan (Miller & Valasek, 2015; Kang et al., 2016; Janbi, 2025). Oleh karena itu, diperlukan mekanisme keamanan tambahan berupa *Intrusion Detection System* (IDS) yang mampu mendeteksi aktivitas mencurigakan secara akurat dan efisien. Pendekatan berbasis machine learning telah banyak digunakan dalam *Intrusion Detection System* (IDS) karena kemampuannya dalam mengenali pola serangan yang kompleks, beradaptasi terhadap karakteristik data jaringan yang dinamis, serta memberikan performa deteksi yang lebih baik dibandingkan metode berbasis aturan (Sommer & Paxson, 2010; Buczak & Guven, 2016; Janbi, 2025). Namun, tantangan utama dari pendekatan ini adalah pemilihan fitur yang relevan dan efisiensi model.

Penelitian ini bertujuan untuk mengatasi permasalahan tersebut dengan mengombinasikan metode seleksi fitur Information Gain dan algoritma XGBoost guna membangun sistem deteksi serangan yang akurat, stabil, dan efisien pada lingkungan *Internet of Vehicles*.

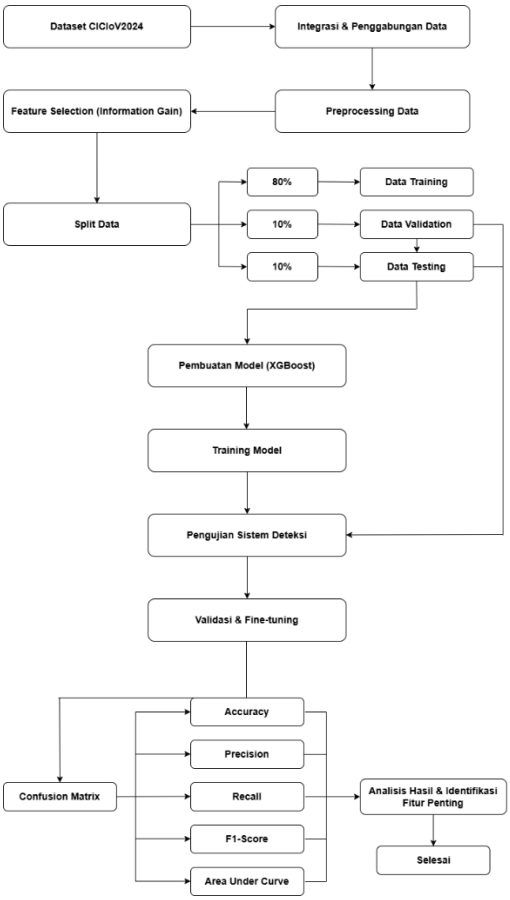
2. KAJIAN TEORITIS

Internet of Vehicles (IoV) merupakan sistem jaringan cerdas yang mengintegrasikan kendaraan, infrastruktur jalan, dan sistem komunikasi untuk mendukung layanan transportasi yang aman dan efisien, namun juga memperluas permukaan serangan siber pada sistem kendaraan (Kaiwartya et al., 2016; Alcaraz et al., 2018; Janbi, 2025). Dalam IoV, keamanan jaringan menjadi aspek krusial karena serangan siber dapat berdampak langsung pada keselamatan pengguna. *Controller Area Network* (CAN Bus) adalah protokol komunikasi utama dalam kendaraan yang memungkinkan pertukaran data antar ECU, namun tidak memiliki mekanisme autentikasi dan enkripsi sehingga rentan terhadap berbagai serangan siber seperti *injection*, *spoofing*, dan DoS (Woo et al., 2015; Miller & Valasek, 2015; Janbi, 2025).

Intrusion Detection System (IDS) berfungsi untuk memantau dan menganalisis lalu lintas jaringan guna mendeteksi aktivitas abnormal. Pendekatan *machine learning* pada IDS memungkinkan sistem belajar dari data historis untuk membedakan antara trafik normal dan serangan. XGBoost merupakan algoritma ensemble berbasis *gradient boosting* yang dikenal memiliki performa tinggi, kemampuan generalisasi yang baik, serta efisiensi komputasi yang unggul dalam menangani data berskala besar dan kompleks, sehingga banyak digunakan dalam tugas klasifikasi keamanan siber (Chen & Guestrin, 2016; Neto et al., 2021; Janbi, 2025). Sementara itu, Information Gain digunakan sebagai metode seleksi fitur untuk mengukur tingkat kepentingan fitur berdasarkan kontribusinya terhadap proses klasifikasi, sehingga membantu meningkatkan akurasi dan efisiensi model *machine learning* dengan mengurangi dimensi data (Quinlan, 1986; Alshamrani et al., 2019; Janbi, 2025).

3. METODE PENELITIAN

Bagian ini menjelaskan dataset yang digunakan, persiapan data, seleksi fitur, penentuan model deteksi, pengujian dan validasi model deteksi. Seluruh digambarkan pada alur penelitian yang disajikan pada gambar 1.



Gambar 1. Alur Penelitian

Secara keseluruhan, proses penelitian ini terdiri atas beberapa tahapan utama, yaitu:

1. Dataset CICIoV2024

   Penelitian ini diawali dengan penggunaan dataset CICIoV2024 yang dikembangkan oleh Canadian Institute for Cybersecurity (CIC). Dataset ini berisi data komunikasi *Controller Area Network* (CAN Bus) pada kendaraan Ford tahun 2019 yang merepresentasikan lalu lintas normal (*benign*) serta berbagai jenis serangan, seperti *Denial of Service* (DoS) dan *spoofing attack*. Dataset disimpan dalam format CSV dan digunakan sebagai sumber utama dalam pengembangan sistem deteksi intrusi berbasis *Internet of Vehicles* (IoV).

2. Integrasi & Penggabungan Data

   Pada tahap ini dilakukan proses integrasi dan penggabungan seluruh data komunikasi CAN yang berasal dari berbagai skenario pengujian ke dalam satu dataset utama. Proses ini bertujuan untuk memastikan konsistensi struktur data serta memudahkan proses analisis dan pemodelan. Selain itu, data juga dikonversi ke format numerik desimal agar kompatibel dengan algoritma pembelajaran mesin.

3. Preprocessing Data

   Tahap preprocessing dilakukan untuk menyiapkan data sebelum digunakan dalam proses pelatihan model. Proses ini meliputi pembersihan data dengan menghapus data duplikat dan menangani nilai kosong, normalisasi untuk menyamakan skala fitur numerik, serta *encoding* label kategorikal menjadi bentuk numerik. Tahap ini sangat penting untuk meningkatkan kualitas data dan kinerja model.

4. *Feature Selection* (*Information Gain*)

   Setelah data diproses, dilakukan seleksi fitur menggunakan metode *Information Gain*. Metode ini mengukur kontribusi setiap fitur terhadap hasil klasifikasi berdasarkan pengurangan nilai entropi. Fitur dengan nilai *Information Gain* tertinggi dipilih karena dianggap paling relevan dalam membedakan antara lalu lintas normal dan serangan, sehingga dapat mengurangi kompleksitas model dan meningkatkan akurasi.

5. Split Data

   Dataset hasil seleksi fitur kemudian dibagi menjadi tiga subset, yaitu 80% data training, 10% data validation, dan 10% data testing. Pembagian ini dilakukan secara acak dan

proporsional untuk memastikan distribusi kelas tetap seimbang serta menghasilkan evaluasi model yang objektif dan representatif.

6. Pembuatan Model (XGBoost)

   Pada tahap ini dilakukan perancangan model klasifikasi menggunakan algoritma *Extreme Gradient Boosting* (XGBoost). Parameter awal seperti *learning rate*, *maximum depth*, jumlah estimator, dan subsample ditentukan sebagai konfigurasi awal model. XGBoost dipilih karena kemampuannya dalam menangani dataset besar, fitur numerik, serta memiliki mekanisme regularisasi untuk mengurangi risiko *overfitting*.

7. *Training* Model

   Model XGBoost dilatih menggunakan data training dengan memanfaatkan data validation untuk memantau kinerja model selama proses pelatihan. Proses *training* dilakukan secara iteratif untuk mempelajari pola lalu lintas CAN dan karakteristik serangan secara optimal, sehingga model mampu membedakan data normal dan serangan dengan baik.

8. Pengujian Sistem Deteksi

   Setelah proses pelatihan selesai, model diuji menggunakan data testing yang belum pernah digunakan sebelumnya. Tahap ini bertujuan untuk mengevaluasi kemampuan model dalam mendeteksi serangan baru serta mengukur tingkat generalisasi model terhadap data yang tidak terlihat selama proses training.

9. Validasi & *Fine-tuning*

   Jika hasil pengujian belum optimal, dilakukan proses validasi dan *fine-tuning* terhadap parameter model. Penyesuaian parameter seperti *learning rate*, *maximum depth*, dan jumlah estimator dilakukan untuk meningkatkan performa model berdasarkan hasil evaluasi pada data *validation*.

10. *Confusion Matrix*

    Hasil klasifikasi model kemudian dianalisis menggunakan *confusion matrix*. *Matriks* ini menampilkan distribusi prediksi yang benar dan salah pada setiap kelas, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). *Confusion matrix* menjadi dasar dalam memahami kesalahan klasifikasi yang dilakukan oleh model.

11. Evaluasi Model

    Berdasarkan confusion matrix, dilakukan perhitungan metrik evaluasi utama, yaitu *accuracy, precision, recall, F1-score*, dan *Area Under Curve* (AUC). Metrik-metrik ini digunakan untuk mengukur efektivitas dan keandalan model dalam mendeteksi serangan pada jaringan kendaraan berbasis IoV.

12. Analisis Hasil & Identifikasi Fitur Penting

    Pada tahap ini dilakukan analisis mendalam terhadap hasil evaluasi model serta identifikasi fitur CAN yang paling berpengaruh berdasarkan nilai *feature importance* dari XGBoost. Analisis ini membantu memahami karakteristik serangan dan kontribusi setiap fitur terhadap keputusan model.

13. Selesai

    Tahap akhir penelitian adalah penarikan kesimpulan berdasarkan hasil yang diperoleh serta evaluasi efektivitas metode *Information Gain*–XGBoost dalam meningkatkan akurasi sistem deteksi intrusi pada jaringan kendaraan berbasis *Internet of Vehicles*.

4. HASIL DAN PEMBAHASAN

Pengumpulan Data dan Lingkup Penelitian

Penelitian ini menggunakan dataset CICIoV2024 yang dikembangkan oleh Canadian Institute for Cybersecurity (CIC). Data dikumpulkan dari kendaraan Ford 2019 melalui komunikasi *Controller Area Network* (CAN Bus) dengan skenario *trafik* normal dan serangan siber, khususnya *Denial of Service* (DoS) dan *spoofing attack*. Dataset tersebut merepresentasikan kondisi komunikasi kendaraan nyata dalam lingkungan *Internet of Vehicles* (IoV). Proses penelitian dilakukan secara offline menggunakan data sekunder tanpa melibatkan pengambilan data langsung di lapangan.

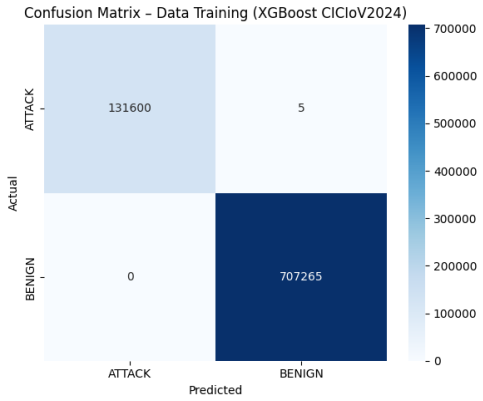
Confusion Matrix

Untuk mengevaluasi kinerja model klasifikasi secara menyeluruh, digunakan metode confusion matrix sebagai alat utama dalam analisis hasil prediksi. *Confusion matrix* memberikan informasi rinci mengenai kesesuaian antara label aktual dan hasil prediksi model pada setiap kelas, sehingga memungkinkan identifikasi prediksi yang benar maupun kesalahan klasifikasi. Analisis ini dilakukan pada tiga jenis data, yaitu data training, validation, dan testing, dengan tujuan untuk menilai kemampuan pembelajaran model, mengukur tingkat

generalisasi terhadap data baru, serta memastikan bahwa model tidak mengalami overfitting atau underfitting. Pembahasan berikut akan menguraikan hasil confusion matrix pada masing-masing jenis data tersebut secara sistematis.

Confusion Matrix Data Training

Evaluasi pada data training menggunakan *confusion matrix* untuk menilai kemampuan model dalam mempelajari pola data latih berdasarkan nilai *TP*, *TN*, *FP*, dan *FN*. Hasil evaluasi ini memberikan gambaran awal performa model dan disajikan pada gambar 2.

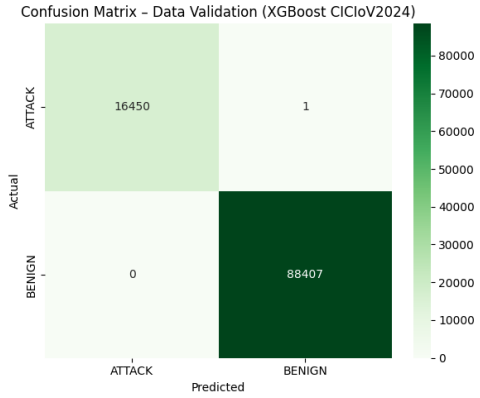


Gambar 2. *Confusion Matrix* - Data *Training*

Berdasarkan gambar 2 tersebut menunjukkan bahwa model XGBoost bekerja sangat optimal dalam membedakan antara kelas ATTACK dan BENIGN. Dari 131600 data serangan (ATTACK), model berhasil mengklasifikasikan semuanya dengan benar kecuali 5 data yang salah diprediksi sebagai BENIGN. Sementara itu, seluruh 707265 data BENIGN berhasil diprediksi dengan benar tanpa adanya kesalahan klasifikasi. Pola ini menggambarkan bahwa model memiliki kemampuan deteksi yang sangat kuat terhadap lalu lintas normal maupun serangan, dengan tingkat kesalahan yang hampir tidak ada.

Confusion Matrix Data Validation

Confusion matrix pada data validation digunakan untuk mengevaluasi kemampuan generalisasi model terhadap data yang tidak digunakan dalam pelatihan. Perbandingan nilai TP, TN, FP, dan FN dengan data training menjadi indikator terjadinya overfitting atau underfitting. Hasil evaluasi disajikan pada gambar 3.

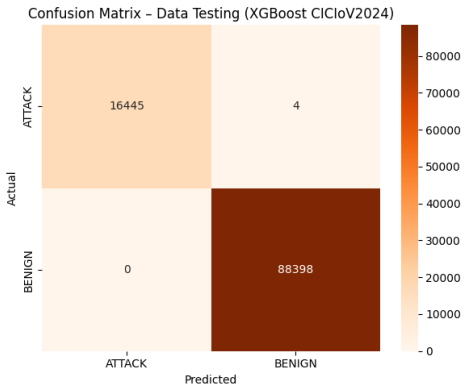


Gambar 3. *Confusion Matrix* - Data Validation

Berdasarkan pada gambar 3 tersebut menunjukkan performa model XGBoost dalam melakukan klasifikasi pada data validasi dataset CICIoV2024. Dari grafik terlihat bahwa model mampu mengidentifikasi hampir seluruh sampel dengan benar. Untuk kelas ATTACK, terdapat 16.450 sampel yang berhasil diprediksi dengan benar (True Positive), sementara hanya 1 sampel yang salah diklasifikasikan sebagai BENIGN (False Negative). Untuk kelas BENIGN, seluruh 88.407 sampel berhasil diklasifikasikan dengan benar (True Negative) tanpa ada satupun yang salah diprediksi sebagai ATTACK (False Positive = 0). Matriks ini menunjukkan bahwa model memiliki kemampuan deteksi serangan yang sangat tinggi sekaligus minim kesalahan dalam mengidentifikasi trafik yang benign. Secara keseluruhan, performa model sangat kuat dan stabil, dengan tingkat accuracy, precision, dan recall yang mendekati sempurna pada data validasi tersebut.

Confusion Matrix Data Testing

Evaluasi akhir menggunakan confusion matrix pada data testing untuk menilai kinerja model pada data yang benar-benar baru. Nilai TP, TN, FP, dan FN digunakan untuk menghitung metrik evaluasi utama, dan hasilnya disajikan pada gambar 4.



Gambar 4. Confusion Matrix - Data Testing

Berdasarkan pada gambar 4.7 tersebut menunjukkan bahwa model XGBoost tetap mempertahankan performa yang sangat baik ketika diuji pada data yang benar-benar terpisah dari proses training dan validasi. Dari 16449 data berlabel ATTACK, sebanyak 16445 berhasil diprediksi dengan benar, sementara hanya 4 data yang keliru diklasifikasikan sebagai BENIGN. Di sisi lain, seluruh 88398 data BENIGN berhasil diprediksi dengan tepat tanpa adanya kesalahan. Hasil ini memperlihatkan bahwa model memiliki kemampuan generalisasi yang kuat dengan tingkat false negative yang sangat rendah dan tanpa false positive. Dengan pola prediksi yang hampir sempurna, model terbukti andal dalam membedakan lalu lintas normal dan serangan pada tahap pengujian.

Hasil Analisis Performa Model

Analisis metrik dilakukan untuk memahami performa model pada setiap jenis data yang terdiri dari data *training, validation, dan testing*. Evaluasi ini bertujuan untuk melihat tingkat generalisasi model serta memastikan tidak terjadi *overfitting*. Metrik yang dianalisis meliputi *accuracy, precision, recall, F1-Score*, serta *TPR, FPR, TNR,* dan *FNR* per kelas. Berikut hasil analisi metrix akan disajikan dalam bentuk tabel berikut.

Performa pada Training Set

Hasil evaluasi *training* menunjukkan bagaimana model mempelajari pola dari data. Nilai metrik ditampilkan pada Tabel 1.

Tabel 1. Hasil Evaluasi Training

Metrik	Nilai (%)
Accuracy	99,9994%
Precision	100%
Recall	99,9962%
F1-Score	99,9981%

Performa pada Validation Set

Tahap validasi digunakan untuk menilai stabilitas model saat mengenali data yang tidak digunakan pada pelatihan. Nilai metrik ditampilkan pada tabel 2.

Tabel 2. Hasil Metrik Evaluasi Validation

Metrik	Nilai (%)
Accuracy	99,9990%
Precision	100%
Recall	99,9939%
F1-Score	99,9969%

Performa pada Testing Set

Testing set memberikan gambaran performa model terhadap data baru yang benar-benar tidak terlihat sebelumnya. Nilai metrik ditampilkan pada tabel 3.

Tabel 3. Hasil Metrik Evaluasi Testing

Metrik	Nilai (%)
Accuracy	99,9962%
Precision	100%
Recall	99,9757%
F1-Score	99,9878%

Analisis Pengaruh *Feature Selection*

Penerapan *Information Gain* terbukti efektif dalam mengidentifikasi fitur-fitur yang paling berkontribusi terhadap proses klasifikasi. Pengurangan jumlah fitur tidak menurunkan performa model secara signifikan, justru meningkatkan efisiensi komputasi dan mengurangi risiko *overfitting*. Hasil ini menunjukkan bahwa tidak semua fitur CAN Bus memiliki kontribusi yang sama dalam mendeteksi serangan, sehingga seleksi fitur menjadi tahap penting dalam pembangunan sistem IDS pada IoV.

Keterkaitan dengan Konsep Dasar dan Penelitian Sebelumnya

Hasil penelitian ini selaras dengan konsep dasar *Intrusion Detection System* berbasis machine learning, yang menyatakan bahwa model pembelajaran mampu mengenali pola serangan yang kompleks pada data jaringan. Temuan ini juga mendukung penelitian sebelumnya yang menyebutkan bahwa algoritma XGBoost memiliki performa unggul dalam klasifikasi data keamanan siber, serta bahwa seleksi fitur berbasis *Information Gain* dapat meningkatkan efisiensi dan akurasi model. Dengan demikian, hasil penelitian ini memperkuat temuan-temuan sebelumnya dalam konteks keamanan jaringan kendaraan.

Implikasi Hasil Penelitian

Secara teoritis, penelitian ini memberikan kontribusi dalam pengembangan pendekatan IDS berbasis *feature selection* dan *ensemble learning* pada lingkungan IoV. Secara terapan, hasil penelitian ini berpotensi diimplementasikan sebagai sistem pendukung keamanan komunikasi CAN Bus untuk mendeteksi serangan DoS dan spoofing secara lebih efektif dan efisien pada kendaraan modern.

5. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa algoritma XGBoost yang dikombinasikan dengan *feature selection Information Gain* mampu mendeteksi serangan DoS dan *spoofing* pada jaringan *Internet of Vehicles* dengan tingkat akurasi yang sangat tinggi. Seleksi fitur menggunakan *Information Gain* terbukti meningkatkan efisiensi dan kemampuan generalisasi model dengan tetap mempertahankan performa optimal. Dengan demikian, kombinasi metode ini sangat potensial untuk diterapkan sebagai *Intrusion Detection System* (IDS) pada lingkungan IoV.

Sebagai saran, penelitian selanjutnya dapat mengembangkan pendekatan ini dengan menambahkan jenis serangan lain, menggunakan dataset yang lebih beragam, atau mengombinasikan XGBoost dengan metode *deep learning* untuk meningkatkan kemampuan deteksi pada skenario yang lebih kompleks.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Dinamika Bangsa serta seluruh pihak yang telah memberikan dukungan dan fasilitas dalam pelaksanaan penelitian ini. Artikel ini merupakan bagian dari Tugas Akhir Program Studi Teknik Informatika.

DAFTAR REFERENSI

N. F. Janbi, “AI-Driven Intrusion Detection in IoV Communication: Insights from CICIoV2024 Dataset,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 3, pp. 272–282, 2025, doi: 10.14569/IJACSA.2025.0160327.

M. S. Korium, M. Saber, A. Beattie, A. Narayanan, S. Sahoo, and P. H. J. Nardelli, “Intrusion detection system for cyberattacks in the Internet of Vehicles environment,” *Ad Hoc Networks*, vol. 153, no. November 2023, 2024, doi: 10.1016/j.adhoc.2023.103330.

L. Yang, A. Shami, G. Stevens, and S. De Rusett, “LCCDE: A Decision-Based Ensemble Framework for Intrusion Detection in the Internet of Vehicles,” *Proc. - IEEE Glob. Commun. Conf. GLOBECOM*, pp. 3545–3550, 2022, doi: 10.1109/GLOBECOM48099.2022.10001280.

M. Alharthi, F. Medjek, and D. Djenouri, “Ensemble Learning Approaches for Multi-Class Intrusion Detection Systems for the Internet of Vehicles (IoV): A Comprehensive Survey,” *Futur. Internet*, vol. 17, no. 7, pp. 1–42, 2025, doi: 10.3390/fi17070317.

M. A. Uddin, N. H. Chu, R. Rafeh, and M. Barika, “A Scalable Hierarchical Intrusion Detection System for Internet of Vehicles,” *IEEE Internet Things J.*, vol. XX, no. Xx, pp. 1–15, 2025, doi: 10.1109/JIOT.2025.3590966.

K. A. Binsaeed and A. M. Hafez, “Enhancing Intrusion Detection Systems with XGBoost Feature Selection and Deep Learning Approaches,” *Int. J. Adv. Comput. Sci. Appl.*, vol.

14, no. 5, pp. 1084–1098, 2023, doi: 10.14569/IJACSA.2023.01405112.

S. S. Dhaliwal, A. Al Nahid, and R. Abbas, "Effective intrusion detection system using XGBoost," *Inf.*, vol. 9, no. 7, pp. 1–28, 2018, doi: 10.3390/info9070149.

Y. Hosain and M. Çakmak, "XAI-XGBoost: an innovative explainable intrusion detection approach for securing internet of medical things systems," *Sci. Rep.*, vol. 15, no. 1, pp. 1–17, 2025, doi: 10.1038/s41598-025-07790-0.

M. Nazeer, A. Alasiry, M. Qayyum, V. K. Madhan, G. Patil, and P. Srilatha, "Enhancing Cyber Security in Autonomous Vehicles: A Hybrid XG Boost-Deep Learning Approach for Intrusion Detection in the CAN Bus," *J. Eur. des Syst. Autom.*, vol. 57, no. 5, pp. 1295–1304, 2024, doi: 10.18280/jesa.570505.

Y. Dong, K. Chen, Y. Peng, and Z. Ma, "Comparative Study on Supervised versus Semi-supervised Machine Learning for Anomaly Detection of In-vehicle CAN Network," *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, vol. 2022-Octob, pp. 2914–2919, 2022, doi: 10.1109/ITSC55140.2022.9922235.

E. Alalwany, I. Mahgoub, B. Alsharif, and A. Alfahaid, "An Intelligent Ensemble-Based Detection of In-Vehicle Network Intrusion," *Appl. Sci.*, vol. 15, no. 12, pp. 1–22, 2025, doi: 10.3390/app15126869.

J. Sun, G. Yang, Y. Chen, H. Wu, and X. Liu, "Energy-Balance-Based Out-of-Distribution Detection of Skin Lesions," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 2, pp. 535–544, 2025, doi: 10.14569/IJACSA.2025.0160255.

T. A. Alhaj, M. M. Siraj, A. Zainal, H. T. Elshoush, and F. Elhaj, "Feature selection using information gain for improved structural-based alert correlation," *PLoS One*, vol. 11, no. 11, pp. 1–18, 2016, doi: 10.1371/journal.pone.0166017.

E. C. P. Neto *et al.*, "CICIoV2024: Advancing realistic IDS approaches against DoS and spoofing attack in IoV CAN bus," *Internet of Things (Netherlands)*, vol. 26, 2024, doi: 10.1016/j.iot.2024.101209.

G. Yanginlar, "Internet of Things (IoT) in Intelligent Transportation Systems: Benefits and Challenges of Implementation," *Eurasia Proc. Sci. Technol. Eng. Math.*, vol. 27, no. June, pp. 16–23, 2024, doi: 10.55549/epstem.1517792.

R. A. Khalil, Z. Safelnasr, N. Yemane, M. Kedir, A. Shafiqurrahman, and N. Saeed, "Advanced Learning Technologies for Intelligent Transportation Systems: Prospects and Challenges," *IEEE Open J. Veh. Technol.*, vol. 5, no. March, pp. 397–427, 2024, doi: 10.1109/OJVT.2024.3369691.

B. A. Tanaji and S. Roychowdhury, "A Survey of Cybersecurity Challenges and Mitigation Techniques for Connected and Autonomous Vehicles," *IEEE Trans. Intell. Veh.*, pp. 1–18, 2024, doi: 10.1109/tiv.2024.3493938.

S. Khanam, I. Bin Ahmedy, M. Y. Idna Idris, M. H. Jaward, and A. Q. Bin Md Sabri, "A Survey of Security Challenges, Attacks Taxonomy and Advanced Countermeasures in the Internet of Things," *IEEE Access*, vol. 8, pp. 219709–219743, 2020, doi: 10.1109/ACCESS.2020.3037359.

C. Zhang, N. Wang, Y. T. Hou, and W. Lou, "Machine Learning-Based Intrusion Detection Systems: Capabilities, Methodologies, and Open Research Challenges," *Authorea Prepr.*, no. Ml, 2025, doi: 10.36227/techrxiv.173627464.48290242/v1.

O. Odukha, "Is the Automotive Industry Ready for the Internet of Vehicles ?," pp. 1–11, 2023.

M. Althunayyan, A. Javed, and O. Rana, “A Survey of Learning-Based Intrusion Detection Systems for In-Vehicle Network,” no. 2019, 2025, [Online]. Available: http://arxiv.org/abs/2505.11551

S. Muneer, U. Farooq, A. Athar, M. A. Raza, T. M. Ghazal, and S. Sakib, “A Critical Review of Artificial Intelligence Based Approaches in Intrusion Detection: A Comprehensive Analysis,” *J. Eng. (United Kingdom)*, vol. 2024, 2024, doi: 10.1155/2024/3909173.

A. S. Mirkhail and Z. Xinyou, “Deep Learning for Anomaly Detection in IoT Healthcare Systems,” *Int. Res. J. Multidiscip. Scope*, vol. 6, no. 2, pp. 1480–1494, 2025, doi: 10.47857/irjms.2025.v06i02.03768.

M. N. Vieira, L. P. Oliveira, and L. Carneiro, “A Comparative Analysis of Machine Learning Algorithms for Distributed Intrusion Detection in IoT Networks,” *Lect. Notes Networks Syst.*, vol. 449 LNNS, pp. 249–258, 2022, doi: 10.1007/978-3-030-99584-3_22.

I. Can, “CIC IoV dataset 2024 Advancing realistic IDS approaches against DoS and spoofing attack in IoV,” pp. 1–8, 2024.

B. Karthiga, D. Durairaj, N. Nawaz, T. K. Venkatasamy, G. Ramasamy, and A. Hariharasudan, “Intelligent Intrusion Detection System for VANET Using Machine Learning and Deep Learning Approaches,” *Wirel. Commun. Mob. Comput.*, vol. 2022, no. October, 2022, doi: 10.1155/2022/5069104.