Deteksi Serangan berbasis Machine Learning pada Internet of Vehicle

Fauzi Adi Rafrastara[1], Wildanil Ghozi[1*], Agung Wardoyo[2]
[1]Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang
[2]Program Studi Rekam Medis, Fakultas Kesehatan, Universitas Dian Nuswantoro, Semarang

Email: wildanil.ghozi@dsn.dinus.ac.id

Abstract

The Internet of Things (IoT) technology has become commonplace in the daily lives of people worldwide. It is not only applied in homes, offices, or urban areas but is also increasingly found in vehicles, including cars. IoT applied to vehicles is commonly referred to the Internet of Vehicles (IoV). The widespread adoption of IoV is due to its ability to enhance user comfort, safety, and efficiency when driving. Some implementations of IoV include vehicle health monitoring and smart parking. On the other hand, the involvement of internet technology in vehicle operations actually opens up security gaps which can certainly have a negative impact on users, vehicles and the surrounding environment. Cyberattacks such as Denial of Service (DoS) and spoofing pose risks by disrupting the sensors within vehicles and interfering communication among involved devices. Therefore, detecting cyberattacks in IoV is essential, and one effective approach is to involve machine learning. In this research experiment, we evaluated the performance of three machine learning algorithms in detecting IoV attacks, including determining whether the attacks are DoS or spoofing. We used a publicly available dataset called CICIoV2024, created in 2024. The freshness of the dataset is crucial in security-related research, given that cyberattack patterns continually evolve over time. The three algorithms used in this study were Decision Tree, Naive Bayey, and Logistic Regression. The Naive Bayes algorithm performed the best, achieving 98.10% of accuracy and a 98.00 of F1-Score.

Keywords: Internet of Things, Internet of Vehicles, Cyber Attack, Attack Detection, Machine Learning

Abstrak

Teknologi Internet of Things (IoT) sudah jamak dimanfaatkan dalam keseharian masyarakat dunia. Tidak hanya diaplikasikan di rumah, kantor atau wilayah kota, pemanfaatkan IoT juga mulai banyak dijumpai pada kendaraan, termasuk mobil. IoT yang diterapkan pada kendaraan, umumnya disebut dengan Internet of Vehicles (IoV). Maraknya penggunaan IoV disebabkan karena teknologi tersebut dapat meningkatkan kenyamanan, keamanan dan efisiensi pengguna dalam berkendara. Beberapa implementasi IoV, diantaranya yaitu untuk memantau kesehatan kendaraan dan juga untuk smart parking. Di sisi lain, pelibatan teknologi internet dalam operasional kendaraan justru membuka celah keamanan yang tentunya dapat berdampak buruk bagi pengguna, kendaraan, dan lingkungan sekitar. Serangan siber seperti Denial of Service (DoS) dan Spoofing beresiko melumpuhkan sensor-sensor yang ada pada kendaraan, sekaligus mengacaukan komunikasi pada perangkat-perangkat yang terlibat. Oleh karena itu, mendeteksi serangan siber pada IoV menjadi sangat esensial, dan salah satu caranya yaitu dengan melibatkan machine learning. Eksperimen pada penelitian ini dilakukan guna mengevaluasi performa tiga algoritma machine learning dalam mendeteksi ada tidaknya serangan pada IoV, termasuk menentukan apakah serangan yang terjadi adalah berupa DoS atau Spoofing. Dataset yang digunakan merupakan dataset publik bernama CICIoV2024 yang dibuat pada tahun 2024. Kebaruan dataset memiliki peran yang sangat penting dalam suatu penelitian terkait dunia keamanan mengingat pola serangan siber terus berevolusi dari waktu ke waktu. Tiga algoritma yang digunakan pada

penelitian ini adalah Decision Tree, Naive Bayes, dan Logistic Regression. Hasilnya, algoritma Naive Bayes mendapatkan performa terbaik, yaitu 98.10% untuk akurasi dan 98.00 untuk F1-Score.

1. Pendahuluan

Internet of Things (IoT) telah menjadi bagian tak terpisahkan dari kehidupan modern [4]. Teknologi ini tidak hanya diaplikasikan di rumah, kantor, atau wilayah kota, tetapi juga merambah ke sektor transportasi, khususnya pada kendaraan. Kombinasi antara teknologi IoT dan kendaraan dikenal dengan istilah Internet of Vehicles (IoV) [21]. IoV merupakan hasil integrasi dan inovasi antara internet, industri elektronika, industri automobile dan industri-industri lain guna menciptakan sistem transportasi yang green dan efisien [20].

IoV merupakan konsep teknologi yang menghadirkan komunikasi antar kendaraan atau vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-everything (V2X), vehicle-to-ecosystem (V2E), serta vehicle-to-surroundings (V2S) [10]. Ragam komunikasi tersebut pada akhirnya akan menghasilkan pertukaran informasi yang berguna untuk kebutuhan *traffic management*, *driver assistance*, dan *safety improvements* [6]. Oleh karena itu, tidak mengherankan apabila IoV disebut sebagai salah satu pemain utama dalam perkembangan teknologi *Intelligent Transportation System (ITS)* di masa depan [8],[3].

Implementasi IoV menawarkan berbagai manfaat bagi pengguna kendaraan, khususnya dalam hal kenyamanan, keamanan dan efisiensi dalam berkendara [13]. Beberapa implementasi IoV diantaranya yaitu untuk memantau kesehatan kendaraan dan juga untuk smart parking, mulai dari mendeteksi area parkir yang tersedia hingga membuat mobil melakukan parkir secara mandiri [9],[5]. Selain itu, IoV juga dapat digunakan pada perusahaan pengiriman untuk mengoptimalkan rute perjalanan, melacak pengiriman secara real-time, serta memberikan estimasi yang akurat kepada user terkait pengiriman barangnya [6].

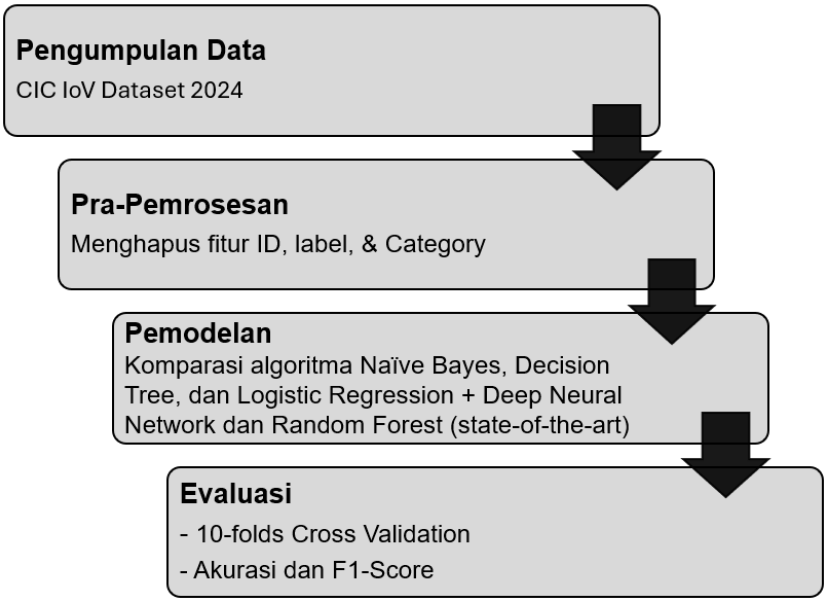
Meskipun IoV menghadirkan banyak manfaat, namun adanya keterlibatan teknologi internet memungkinkan munculnya celah keamanan yang dapat beresiko tinggi. Serangan-serangan yang menarget sistem IoV mulai bermunculan dan terus mengalami penigkatan, mulai dari serangan dalam bentuk Denial of Service (DoS) hingga Spoofing [12]. Oleh karena itu, dibutuhkan metode yang dapat secara efektif mendeteksi serangan-serangan tersebut sehingga sistem IoV tetap aman dan terhindar dari kegagalan yang dapat berdampak fatal pada keselamatan pengguna, kendaraan, maupun lingkungan sekitar.

Pada penelitian yang dilakukan oleh [12], terdapat 4 algoritma yang diukur performanya, yaitu Logistic Regression, AdaBoost, DeepNeuralNetwork, dan Random Forest. Ke-empat algoritma tersebut diuji pada 2 jenis dataset, yaitu Binary Dataset dan Decimal. Hasilnya, Deep Neural Network memperoleh skor tertinggi, baik pada Binary Dataset maupun Decimal Dataset. Pada Binary Dataset, Deep Neural Network memiliki skor 0.95 untuk akurasi dan 0.63 untuk F1-Score. Sedangkan pada Decimal Dataset, Deep Neural Network mendapatkan skor sebesar 0.96 untuk akurasi dan 0.78 untuk F1-Score. Sayangnya, penelitian ini tidak menggunakan metode Cross Validation dalam membagi data training dan testing, sehingga rawan terjadi overfitting. Selain itu, penelitian ini juga cenderung menggunakan algoritma-algoritma yang kompleks sehingga menambah beban komputasi dimana hasilnya belum tentu lebih baik daripada algoritma-algoritma yang lebih sederhana, seperti Naïve Bayes dan Decision Tree.

Penelitian terkait deteksi serangan siber berbasis machine learning pada IoV juga dilakukan oleh [1]. Pada penelitian ini, peneliti menggunakan dataset CAN-intrusion-dataset yang memuat tiga jenis serangan, yaitu DoS Attack, Fuzzy Attack, dan Attack Free State. Peneliti mengusulkan penggunaan VGG-16, yaitu algoritma berbasis Deep Learning untuk mendeteksi serangan. Hasilnya, VGG-16 berhasil mengungguli 5 algoritma lain seperti kNN, Random Forest, Gradient Boosting, AdaBoost, dan SVM, dengan nilai akurasi sebesar 96%. Di sini, metode pemisahan data training dan testing yang digunakan adalah Split Validation, dengan rasio 70% untuk training dan 30% untuk testing. Tanpa menggunakan metode Cross-Validation, maka akan meningkatkan potensi terjadinya overfitting.

2. Metode

Tahapan yang dilakukan pada penelitian ini meliputi pengumpulan data, pre-processing, pemodelan, dan evaluasi, sebagaimana dapat dilihat pada Gambar 1.



Gambar 1. Tahapan penelitian yang digunakan.

a. Hardware dan Software

Penggunaan perangkat keras (*hardware*) dan perangkat lunak (*software*) yang tepat dapat menunjang kelancaran dan keberhasilan suatu penelitian. Hardware dengan spesifikasi tinggi tanpa didukung oleh software yang tepat, akan menjadi sia-sia. Sebaliknya, software yang tepat tanpa didukung hardware dengan spefisikasi yang sesuai, akan menjadi tidak optimal [15]. Oleh karena itu, pemilihan hardware dan software menjadi sangat penting, terutama dalam penelitian ini, mengingat pemrosesan yang dilakukan adalah pada dataset dengan jumlah lebih dari 1 juta *instances*.

b. Pengumpulan Data

Penelitian ini menggunakan dataset publik yang dikembangkan pada tahun 2024 oleh para peneliti dari University of New Brunswick, Kanada, bernama “CIC IoV Dataset 2024” [19]. Kebaruan dataset sangat penting pada penelitian di bidang *information security*, mengingat pola-pola serangan di dunia siber terus berevolusi dari waktu ke waktu. Detail dari dataset yang digunakan dapat dilihat pada tabel 1.

Tabel 1. Detail dataset yang digunakan.

Nama Dataset	CIC IoV Dataset 2024
Tahun Pembuatan	2024
Jumlah Fitur	11
Jumlah Instances	1.408.249
Jumlah Kelas	6 (Benign, DoS, Gas-Spoofing, Steering Wheel-Spoofing, Speed-Spoofing, dan RPM-Spoofing)

Data IoV yang dihasilkan oleh kendaraan dapat dikelompokkan ke dalam dua jenis, yaitu On-board data dan On-road Data [17]. On-board data merupakan data yang digunakan untuk memonitor status perangkat pada kendaraan, seperti rem, gas, steering wheel, velocity, dan

parameter-paramer mesin lain. Sedangkan On-road data yaitu data tentang data-data yang dihasilkan terhadap lingkungan sekitar selama kendaraan melaju, seperti jarak antar kendaraan, jarak kendaraan dengan objek lain, dan traffic light. CIC IoV Dataset 2024 menggunakan On-board data, dan mayoritas serangannya pun juga terkait dengan spoofing pada data-data tersebut, seperti Gas-Spoofing, Steering Wheel-Spoofing, Speed-Spoofing, dan RPM-Spoofing.

c. Pra-Pemrosesan

Dalam penelitian data mining, pra-pemrosesan atau *pre-processing* merupakan tahapan penyiapan data sebelum pemodelan dengan algoritma *machine learning* dilakukan. Pada tahap ini, yang dilakukan oleh peneliti adalah menghapus beberapa fitur yang tidak berguna, seperti ID, label, dan category. Dengan demikian, tersisa 8 fitur yang siap diproses lebih lanjut. Nama-nama 8 fitur tersebut adalah DATA_0, DATA_1, DATA_2, DATA_3, DATA_4, DATA_5, DATA_6, dna DATA_7.

Panjang payload yang dibawa oleh setiap dataframe berukuran 8 bytes. Pada dataset tersebut, setiap byte dataframe diekstrak menjadi 1 fitur. Dengan demikian, makna fitur DATA_0 yaitu isi byte ke 0 pada data yang ditransmisikan. Demikian juga Data_1, merupakan byte ke 1 pada data yang ditransmisikan. Begitu seterusnya hingga fitur DATA_7, yang berisi byte ke 7 pada data yang ditransmisikan.

d. Pemodelan

Pada penelitian ini, tiga algoritma klasifikasi akan dibandingkan untuk mengetahui performa masing-masing algoritma dan mendapatkan hasil yang terbaik. Ketiga algoritma tersebut yaitu Naïve Bayes, Decision Tree, dan Logistic Regression.

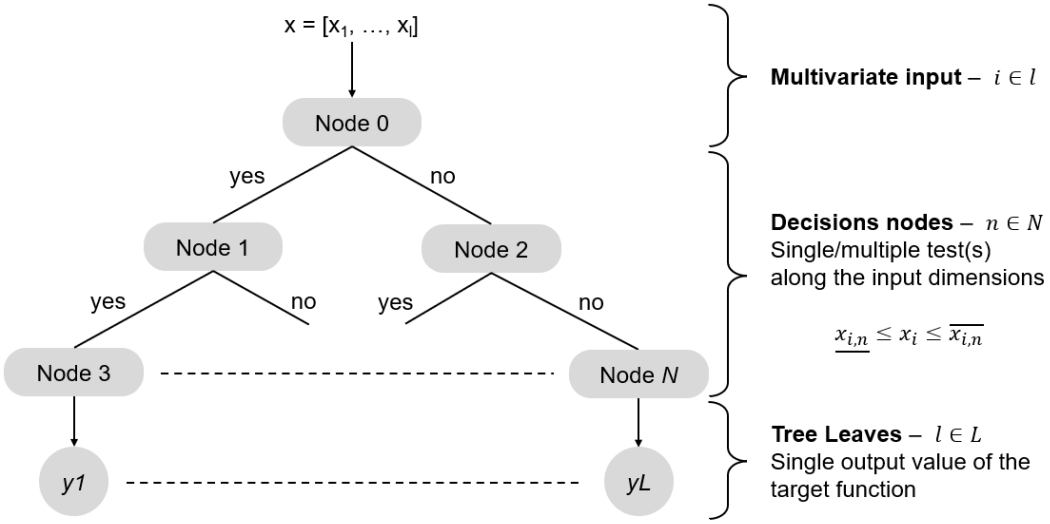
Algoritma Naïve Bayes merupakan salah satu metode klasifikasi yang popular dalam data science, yang dikenal karena kesederhanaan dan efektifitasnya. Algoritma Naïve Bayes bekerja berdasarkan teorema Bayes, dimana algoritma ini memprediksi probabilitas suatu data masuk ke kelas tertentu dengan mempertimbangkan kemunculan atribut-atributnya dalam kelas tersebut. Naïve Bayes dikatakan naïve karena mengasumsikan adanya independensi antar atribut. Meskipun begitu, algoritma Naïve Bayes dapat menghasilkan performa yang kompetitif, mudah dalam implementasinya, dan menghasilkan komputasi yang efisien. Formula yang digunakan pada algoritma Naïve Bayes, dapat dilihat pada *equation* 1 [11],[7].

$$\mathrm{P(C|X)} = \big(\mathrm{P(X|C)} * \mathrm{P(C)}\big)/\mathrm{P(X)} \tag{1}$$

Dimana:

- P(C|X) : probabilitas data X termasuk dalam kelas C.
- P(X|C) : probabilitas kemunculan atribut X dalam kelas C.
- P(C) : probabilitas apriori kelas C.
- P(X) : probabilitas apriori kelas X.

Algoritma yang kedua yaitu algoritma Decision Tree. Algoritma decision tree merupakan algoritma yang dapat digunakan untuk penyelesaian kasus klasifikasi maupun regresi. Algoritma decision tree bekerja dengan memecah data menjadi cabang-cabang berdasarkan atribut tertentu, sehingga mampu memprediksi kelas atau nilai tertentu dengan cara yang mudah dipahami. Terdapat beberapa algoritma klasifikasi yang dikembangkan berbasis decision tree, beberapa diantaranya yaitu: Iterative Dichotomizer 3 (ID3), C4.5, dan CART. Aplikasi Orange juga memiliki pemodelan berbasis *decision tree* dan dapat digunakan melalui widget yang bernama Tree. Widget ini mendukung pemrosesan baik untuk dataset kategorikal maupun numeric. Ilustrasi decision tree dapat dilihat pada gambar 2 [16].



Gambar 2. Ilustrasi decision tree

Algoritma ketiga yaitu Logistic Regression. Algoritma Logistic Regression bekerja dengan memodelkan hubungan antara variabel independent (fitur) dan variabel dependen (kelas/target) menggunakan fungsi logistic [12]. Logistic Regression menghasilkan probabilitas untuk setiap data point, menunjukkan peluangnya untuk masuk ke dalam kelas tertentu. Algoritma Logistic Regression dikenal sebagai algoritma yang sederhana, interpretasi yang mudah, serta kemampuannya menangani data dengan beragam dimensi. Formula yang digunakan pada algoritma Logistic Regression dapat dilihat pada equation 2.

$$P(y \mid x)=1/(1+\exp(-(\beta 0+\sum \beta_i x_i))) \quad (2)$$

Dimana:

- P(y | x) : Probabilitas data point x termasuk dalam kelas y (misalnya, probabilitas suatu traffic dikategorikan sebagai attack).
- β0 : Konstanta intercept (bias).
- β_i : Koefisien regresi untuk variabel independen x_i (menunjukkan pengaruh variabel x_i terhadap probabilitas).
- x_i : Nilai variabel independen x_i (fitur) pada data point.
- exp : Fungsi eksponensial (e).
- Σ : Simbol penjumlahan (menjumlahkan nilai $\beta_i x_i$ untuk semua variabel independen).

e. Evaluasi

Evaluasi merupakan langkah penting dalam data mining untuk menilai performa model yang dihasilkan. Tujuannya dilakukannya evaluasi model adalah untuk memastikan bahwa model tersebut akurat, andal, dan berguna dalam menyelesaikan permasalahan yang dituju. Agar suatu model dapat dievaluasi, maka model perlu divalidasi lebih dulu, baik menggunakan *Split Validation* ataupun *Cross Validation*. Pada penelitian ini, metode yang digunakan adalah *Cross Validation*, dengan k = 10 (*10-folds Cross-Validation*). Dengan menggunakan 10-folds Cross Validation, maka dataset akan dibagi mejadi 10 kelompok, dimana setiap kelompok memiliki ukuran (jumlah data) yang sama. Selanjutnya, pengujian akan dilakukan hingga 10 kali (10-folds).Setiap kelompok akan dilibatkan dalam training sebanyak 9 kali dan testing sebanyak satu kali. Metode ini dipilih karena dapat meminimalisir kasus *overfitting*, jika dibandingkan dengan metode *Split Validation* [2],[14],[18].

Proses berikutnya yaitu mengevaluasi model dengan melibatkan berbagai metrik yang sesuai dengan jenis kasus di data mining. Pada kasus klasifikasi, metrik yang umumnya

digunakan adalah akurasi, recall, presisi, dan F1-Score [18]. Setiap metrik memiliki fungsinya masing-masing. Metrik utama yang digunakan pada penelitian ini adalah akurasi dan F1-Score.

Akurasi merupakan metrik untuk mengukur prediksi yang benar secara keseluruhan. Pada formula Akurasi (Equation 3), TP atau *True Positive* merupakan jumlah berapa kali model memprediksi kelas positive dengan tepat. Sementara itu, TN atau *True Negative* merupakan jumlah berapa kali model memprediksi kelas negative dengan tepat. FP atau *False Positive* merupakan jumlah berapa kali model salah memprediksi kelas Positive. Sedangkan FN atau *False Negative* merupakan jumlah berapa kali model salah memprediksi kelas Negative. Dalam kasus deteksi serangan pada IoV, akurasi berguna untuk memberikan gambaran secara umum tentang performa model dalam mendeteksi suatu traffic, apakah tergolong traffic *benign*, *DoS, Gas-Spoofing, Steering Wheel-Spoofing, Speed-Spoofing, dan RPM-Spoofing*.

$$\text{Akurasi} = \frac{\text{TP+TN}}{(\text{TP+FP+TN+FN})} \tag{3}$$

Metrik kedua yang digunakan adalah F1-Score. F1-Score berguna untuk memberikan gambaran tentang keseimbangan antara presisi (Equation 4) dan recall (Equation 5). Rumus perhitungan F1-Score dapat dilihat pada equation 6. Dalam proses kalkulasinya, F1-Score membutuhkan nilai Presisi dan Recall. Presisi merupakan metrik untuk mengukur proporsi prediksi positif yang benar-benar positif. Di sisi lain, Recall adalah metrik untuk mengukur proporsi data positif yang berhasil diidentifikasi. Kesalahan prediksi terhadap serangan pada IoV, baik positive menjadi negative atau sebaliknya, dapat mengakibatkan dampak yang berbahaya. Oleh karena itu, F1-Score dibutuhkan untuk mendapatkan keseimbangan diantara keduanya.

$$\text{Presisi} = \frac{\text{TP}}{\text{TP + FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \tag{5}$$

$$\text{F1} - \text{Score} = \frac{\text{2 x (Presisi x Recall)}}{\text{Presisi+Recall}} \tag{6}$$

Selanjutnya hasil evaluasi digunakan untuk membandingkan model yang berbeda, memilih model terbaik, dan mengidentifikasi area yang perlu diperbaiki. Evaluasi yang baik akan menghasilkan model data mining yang berkualitas dan memberikan manfaat nyata pada pengembangannya.

3. Hasil dan Pembahasan

Pada dataset yang didownload, mulanya terdapat 12 kolom dan 1.408.249 data (*instances*). Salah satu kolom yang bernama *specific_class* dijadikan sebagai *variabel dependent* (*target*), sementara sisanya sebagai *variabel independent* atau fitur. Pada bagian *target*, terdapat 6 kelas yang akan digunakan pada penelitian ini, yaitu *benign, DoS, Gas Spoofing, Steering Wheel Spoofing*, *Speed Spoofing*, dan *RPM Spoofing*.

Dari 11 fitur yang tersedia, selanjutnya dilakukan langkah pra-pemrosesan dengan menghapus beberapa fitur yang dinilai tidak relevan dengant tujuan penelitian. Pada tahap ini, terdapat 3 fitur yang dihapus, yaitu fitur *ID*, *label*, dan *category*. Fitur ID dihapus karena tidak memiliki korelasi terhadap kelas target. Fitur ini merupakan kode priporitas sekaligus pengenal dari sebuah pesan yang dikirimkan di dalam protocol CAN (Controller Area Network).

Fitur *label* sebetulnya bisa dijadikan sebagai target pada klasifikasi biner karena berisi hasil klasifikasi suatu data *traffic* tergolong *benign* atau *attack*. Namun karena tujuan penelitian ini adalah untuk melakukan klasifikasi secara lebih detail, yaitu ke dalam 6 kelas, maka fitur *label* tidak dibutuhkan. Fitur terakhir yang dihapus yaitu fitur *category*. Fitur ini berisi informasi kategori suatu traffic, apakah tergolong *benign, DoS,* atau *Spoofing*. Sama halnya dengan fitur

label, karena tujuan dari penelitian ini adalah untuk klasifikasi ke dalam 6 kelas yang telah disebutkan sebelumnya, maka fitur ini tidak digunakan. Dengan dihapusnya 3 fitur tersebut, maka kini tersisa 8 fitur yang siap diproses lebih lanjut.

Langkah berikutnya yaitu pemodelan, dengan menerapkan 3 algoritma klasifikasi *machine learning* pada dataset yang telah disiapkan. Ketiga algoritma tersebut yaitu Naïve Bayes, Decision Tree, dan Logistic Regression. Hasil performa ketiga algoritma tersebut akan dibandingkan dengan performa dari 2 algoritma lain yang terdapat pada paper [12], yaitu Deep Neural Network dan Random Forest. Dalam tahapan evaluasi pada penelitian ini, 10-folds Cross Validation digunakan untuk memisahkan data training dan testing sekaligus untuk meminimalisir overfitting, sebelum diproses lebih lanjut dengan mengukur performa akurasi dan F1-Score-nya. Pada penelitian [12], sebetulnya Logistic Regression telah diuji dan menghasilkan performa yang kurang baik, yaitu 0.89 untuk akurasi, serta 0.49 untuk F1-Score. Pada penelitian ini, *Logistic Regression* diuji kembali pada environment yang telah diubah. Sayangnya hasilnya masih belum mampu menandingi beberapa algoritma lain, dan masih berada di bawah angka 0.90, baik untuk akurasi maupun F1-Score. Hasil lengkap pengujian algoritma pada penelitian ini dapat dilihat pada tabel 2.

Tabel 2. Hasil perbandingan antar algoritma.

Nama Algoritma	Akurasi	F1-Score
Naïve Bayes	0.981	0.98
Decision Tree	0.975	0.971
Logistic Regression	0.876	0.842
Deep Neural Network [12]	0.96	0.78
Random Forest [12]	0.96	0.76

Berdasarkan hasil pengujian yang telah dilakukan, algoritma Naïve Bayes berhasil memperoleh skor tertinggi, yaitu 0.981 untuk akurasi, dan 0.98 untuk F1-Score. Algoritma ini juga jauh lebih sederhana dibandingkan dengan Deep Neural Network maupun Random Forest yang digunakan pada paper [12], sehingga berpengaruh pada kompleksitas algoritma. Di antara kelima algoritma yang ada pada tabel 2, performa terendah dimiliki oleh algoritma Logistic Regression. Hal ini dapat dipahami, mengingat algoritma Logistic Regression kurang ideal untuk kasus *multiclass classification*.

4. Kesimpulan

Internet of Vehicles berhasil menghadirkan kenyamanan berbeda bagi para penggunanya. Sayangnya, penggunaan teknologi seperti ini dapat memunculkan masalah baru, khususnya dalam hal keamanan. Dua jenis serangan yang umum terjadi adalah Denial of Service dan Spoofing. Keduanya dapat menimbulkan dampak serius bagi pengguna dan juga kendaraan itu sendiri. Oleh karena itu, metode deteksi serangan pada wilayah IoV menjadi penting untuk terus kembangkan. Penelitian ini dilakukan dengan tujuan untuk mendapatkan algoritma klasifikasi terbaik dalam mendeteksi serangan pada IoV, dengan membandingkan performa algoritma Naïve Bayes, Decision Tree, Logistic Regression, serta 2 algoritma pada penelitian [12], yaitu Deep Neural Network dan Random Forest. Hasilnya, Naïve Bayes memperoleh nilai terbaik, dengan 0.981 untuk akurasi dan 0.98 untuk F1-Score.

5. Referensi

[1]. Ahmed, I., Jeon, G. and Ahmad, A. (2023) 'Deep Learning-Based Intrusion Detection System for Internet of Vehicles', *IEEE Consumer Electronics Magazine*, 12(1), pp. 117–123. Available at: https://doi.org/10.1109/MCE.2021.3139170.

[2]. Battineni, G. *et al.* (2019) 'Comparative Machine-Learning Approach: A Follow-Up Study on Type 2 Diabetes Predictions by Cross-Validation Methods', *Machines*, 7(4), p. 74. Available at: https://doi.org/10.3390/machines7040074.

[3]. Chen, M. *et al.* (2024) 'An attribute-encryption-based cross-chain model in urban internet of vehicles', *Computers and Electrical Engineering*, 115, p. 109136. Available at: https://doi.org/10.1016/j.compeleceng.2024.109136.

[4]. Chung, W. and Cho, T. (2022) 'Complex attack detection scheme using history trajectory in internet of vehicles', *Egyptian Informatics Journal*, 23(3), pp. 499–510. Available at: https://doi.org/10.1016/j.eij.2022.05.002.

[5]. Dheeven, T.A. *et al.* (2024) 'IoT based sensor enabled vehicle parking system', *Measurement: Sensors*, 31, p. 100953. Available at: https://doi.org/10.1016/j.measen.2023.100953.

[6]. Djenouri, Y. *et al.* (2024) 'Enhancing smart road safety with federated learning for Near Crash Detection to advance the development of the Internet of Vehicles', *Engineering Applications of Artificial Intelligence*, 133, p. 108350. Available at: https://doi.org/10.1016/j.engappai.2024.108350.

[7]. Gao, H. (2023) 'Spam sorting based on Naive Bayes Algorithm', in. *2023 International Conference on Blockchain Technology and Applications (ICBTA)*.

[8]. Haodudin Nurkifli, E. and Hwang, T. (2023) 'Provably secure authentication for the internet of vehicles', *Journal of King Saud University - Computer and Information Sciences*, 35(8), p. 101721. Available at: https://doi.org/10.1016/j.jksuci.2023.101721.

[9]. Kaur, G. and Garg, H. (2023) 'A novel algorithm for autonomous parking vehicles using adjustable probabilistic neutrosophic hesitant fuzzy set features', *Expert Systems with Applications*, 226, p. 120101. Available at: https://doi.org/10.1016/j.eswa.2023.120101.

[10]. Korium, M.S. *et al.* (2024) 'Intrusion detection system for cyberattacks in the Internet of Vehicles environment', *Ad Hoc Networks*, 153, p. 103330. Available at: https://doi.org/10.1016/j.adhoc.2023.103330.

[11]. Mandala, S. *et al.* (2022) 'DDoS Detection by Using Information Gain-Naïve Bayes', in *2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*. *2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, Bandung, Indonesia: IEEE, pp. 283–288. Available at: https://doi.org/10.1109/ICICyTA57421.2022.10038054.

[12]. Neto, E.C.P. *et al.* (2024) 'CICIoV2024: Advancing realistic IDS approaches against DoS and spoofing attack in IoV CAN bus', *Internet of Things*, 26, p. 101209. Available at: https://doi.org/10.1016/j.iot.2024.101209.

[13]. Qureshi, K.N. *et al.* (2021) 'Internet of Vehicles: Key Technologies, Network Model, Solutions and Challenges With Future Aspects', *IEEE Transactions on Intelligent Transportation Systems*, 22(3), pp. 1777–1786. Available at: https://doi.org/10.1109/TITS.2020.2994972.

[14]. Rafrastara, F.A. *et al.* (2023) 'Performance Improvement of Random Forest Algorithm for Malware Detection on Imbalanced Dataset using Random Under-Sampling Method', *Jurnal Informatika*, 8(2), pp. 113–118.

[15]. Rafrastara, F.A. *et al.* (2024) 'Performance Comparison of k-Nearest Neighbor Algorithm with Various k Values and Distance Metrics for Malware Detection', 8.

[16]. Rigo-Mariani, R. and Yakub, A. (2024) 'Decision Tree Variations and Online Tuning for Real-Time Control of a Building in a Two-Stage Management Strategy', *Energies*, 17(11), p. 2730. Available at: https://doi.org/10.3390/en17112730.

[17]. Sherazi, H.H.R. *et al.* (2019) 'DDoS attack detection: A key enabler for sustainable communication in internet of vehicles', *Sustainable Computing: Informatics and Systems*, 23, pp. 13–20. Available at: https://doi.org/10.1016/j.suscom.2019.05.002.

[18]. Supriyanto, C. *et al.* (2024) 'Malware Detection Using K-Nearest Neighbor Algorithm and Feature Selection', 8.

[19]. University of New Brunswick (2024) 'CIC IoV dataset 2024: Advancing Realistic IDS Approaches against DoS and Spoofing Attack in IoV CAN bus'. Available at: https://www.unb.ca/cic/datasets/iov-dataset-2024.html (Accessed: 9 June 2024).

[20]. Wang, M. and Wang, S. (2021) 'Communication Technology and Application in Internet of Vehicles', in *2021 IEEE 4th International Conference on Information Systems and Computer Aided Education (ICISCAE)*. *2021 IEEE 4th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China: IEEE, pp. 234–237. Available at: https://doi.org/10.1109/ICISCAE52414.2021.9590660.

[21]. Wei, X. (2024) 'Enhancing road safety in internet of vehicles using deep learning approach for real-time accident prediction and prevention', *International Journal of Intelligent Networks*, 5, pp. 212–223. Available at: https://doi.org/10.1016/j.ijin.2024.05.002.