ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

IMPLEMENTASI METODE *ISOLATION FOREST* UNTUK DETEKSI ANOMALI DALAM DATA JARINGAN

Rafli Ardiansyah[1*], Lukman Sunardi[2] , A. Taqwa Martadinata[3]
[1]Program Studi Informatika, Universitas Bina Insan, Lubuklinggau

e-mail: ***[1]**rafli5564@gmail.com, [2]lukmanmmci@gmail.com, [3] marthadinata95@gmail.com
*Corresponding Author

ABSTRAK

Keamanan jaringan menjadi hal kritis di era digital untuk melindungi data dari ancaman siber yang semakin kompleks. Penelitian ini mengkaji penerapan algoritma Isolation Forest dalam mendeteksi anomali pada data jaringan, dipilih karena kemampuannya memproses data besar secara efisien dan mendeteksi anomali dengan cepat. Tahapan penelitian meliputi pengumpulan data menggunakan Wireshark (mencakup parameter seperti panjang paket, protokol, dan alamat), preprocessing (pembersihan data, normalisasi, seleksi fitur), pelabelan data, dan penerapan algoritma. Evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score menunjukkan akurasi 81% dengan rasio pembagian data latih-uji 80:20. Hasil ini membuktikan efektivitas Isolation Forest dalam mengisolasi anomali sekaligus menjaga efisiensi pada dataset besar. Penelitian ini berkontribusi pada pengembangan deteksi anomali berbasis *machine learning*, memperkaya literatur, serta menjadi panduan bagi praktisi untuk meningkatkan keamanan jaringan di masa depan.

Kata Kunci : Deteksi Anomali, Keamanan Jaringan, *Isolation Forest*, *Machine Learning*, Data Jaringan

ABSTRACT

Network security is a critical aspect in the digital era, essential for protecting data from increasingly complex cyber threats. This study examines the implementation of the Isolation Forest algorithm for detecting anomalies in network data, chosen for its efficiency in processing large-scale data and its ability to quickly identify anomalous patterns. The research stages include data collection using Wireshark (capturing parameters such as packet length, protocol, and source/destination addresses), preprocessing (data cleaning, normalization, feature selection), data labeling, and algorithm application. Evaluation using accuracy, precision, recall, and F1-**score** metrics achieved 81% accuracy with an 80:20 train-test split. These results demonstrate the effectiveness of Isolation Forest in isolating anomalies while maintaining efficiency on large datasets. This research contributes to the advancement of machine learning-based anomaly detection**,** enriches academic literature, and provides a practical guide for improving network security in the future.

***Keywords** : Anomaly Detection, Network Security, Isolation Forest, Machine Learning, Network Data*

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

I. PENDAHULUAN

Keamanan jaringan merupakan salah satu aspek yang sangat penting dalam dunia digital saat ini. Seiring dengan berkembangnya teknologi dan meningkatnya ketergantungan terhadap sistem berbasis jaringan, risiko terhadap serangan siber dan anomali jaringan semakin meningkat. Anomali dalam jaringan dapat menjadi indikasi awal dari serangan berbahaya, seperti peretasan, pencurian data, atau gangguan sistem yang dapat menyebabkan kerugian besar bagi individu maupun organisasi.

Sistem deteksi intrusi (*Intrusion Detection System/IDS*) telah banyak digunakan untuk mengidentifikasi ancaman yang tidak terdeteksi oleh sistem keamanan konvensional. Salah satu metode yang mulai banyak digunakan dalam deteksi anomali adalah algoritma berbasis *machine learning*, salah satunya *Isolation Forest*. *Isolation Forest* adalah algoritma yang dirancang khusus untuk mendeteksi anomali dengan mengisolasi data outlier dari data normal melalui struktur pohon keputusan. Metode ini memiliki keunggulan

dalam mendeteksi anomali secara efisien pada dataset besar dengan waktu komputasi yang lebih cepat dibandingkan metode lainnya.

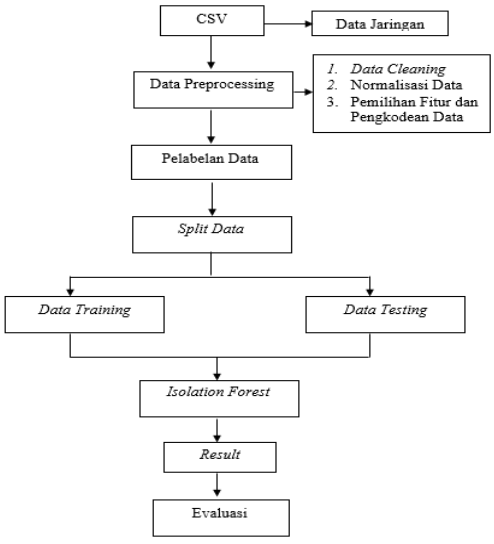
Penelitian ini bertujuan untuk mengimplementasikan metode *Isolation Forest* dalam mendeteksi anomali pada data jaringan. Dengan melakukan analisis dan evaluasi terhadap model yang diterapkan, penelitian ini diharapkan dapat memberikan wawasan lebih lanjut mengenai efektivitas metode ini dalam meningkatkan keamanan jaringan. Evaluasi kinerja model akan dilakukan menggunakan metrik seperti akurasi, *precision*, *recall*, dan *F1-score* untuk menilai seberapa baik algoritma dalam mengidentifikasi anomali.

II. METODOLOGI PENELITIAN

A. METODE PENELITIAN

Penelitian ini memanfaatkan teknik *unsupervised machine learning* untuk deteksi anomali menggunakan metode *Isolation Forest*. Metode ini merupakan bagian dari modul dalam *library machine learning* ***scikit-learn***, yang diimplementasikan menggunakan bahasa pemrograman *Python*. *Unsupervised learning* adalah pendekatan pembelajaran mesin yang bekerja dengan data tanpa label tertentu. Dalam pendekatan ini, mesin secara mandiri mengidentifikasi pola-pola tersembunyi dalam data tanpa memerlukan campur tangan manusia [1].

Dalam penelitian ini, peneliti menerapkan konsep *Machine Learning* dengan mengoptimalkan algoritma *Isolation Forest* untuk mendeteksi anomali pada data jaringan. Data jaringan yang digunakan diperoleh melalui aplikasi *Wireshark*. Terdapat diagram alur konsep penelitian yang menunjukkan tahapan dari pengumpulan data hingga evaluasi model, sebagaimana tertera pada Gambar 1.



Gambar 1. Diagram Konsep Penelitian

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

B. TEKNIK PENGUMPULAN DATA

a) Data Primer

Data primer adalah data yang diperoleh langsung dari sumber yang akan digunakan. Dalam penelitian ini, data primer yang digunakan adalah data jaringan yang diperoleh melalui aplikasi *Wireshark* untuk menangkap paket data yang dikirim dan diterima dalam jaringan. Data ini mencakup informasi seperti panjang paket, protokol, waktu, dan sumber/destinasi. Data tersebut disimpan dalam format *CSV* dan digunakan untuk mendeteksi anomali yang terjadi dalam jaringan, baik yang bersifat normal maupun anomali.

b) Data Sekunder

Data sekunder diperoleh dari sumber-sumber yang sudah ada sebelumnya untuk mendukung penelitian penulis. Data sekunder ini dapat berupa wawasan dari literatur, seperti buku, jurnal, artikel ilmiah, atau informasi dari media sosial yang dapat diakses melalui internet.

C. PENGUJIAN DAN PENGOLAHAN DATA

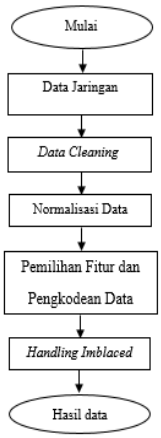
1. Pengujian

Penelitian ini menggunakan *library Confusion Matrix* untuk pengujian pada algoritma *Isolation Forest*. Metode ini diterapkan guna mengevaluasi kinerja dalam mendeteksi anomali. *Confusion Matrix* direpresentasikan sebagai *matriks* berukuran 2x2 yang menyajikan hasil klasifikasi *biner* pada sebuah dataset, terdiri dari empat komponen utama di dalamnya.

2. Pengolahan Data

1) Tahap Preprocessing

Sebelum pengolahan data dilakukan, tahap pertama yang dilaksanakan adalah *Preprocessing* menggunakan bahasa pemrograman *Python*. Tahap ini mencakup beberapa langkah, di antaranya pengambilan data jaringan melalui aplikasi *Wireshark*. Selanjutnya, dilakukan proses Data *Cleaning*, Normalisasi Data, dan Pemilihan Fitur dan Pengkodean Data sebagaimana yang ditunjukkan pada gambar berikut :



Gambar 2. Tahap *Preprocessing*

2) Tahap Labelling

Labelling data merupakan proses menambahkan label pada data untuk mengelompokkannya ke dalam kategori seperti anomali dan normal, yang diperlukan untuk keperluan analisis. Dalam penelitian ini, data jaringan yang digunakan saat ini belum memiliki label, sehingga proses pelabelan akan dilakukan untuk membedakan data normal dari data anomali. Langkah ini sangat penting untuk memungkinkan model deteksi anomali memahami pola dan memberikan hasil yang lebih akurat.

3) Tahap Visualisasi Data

Setelah proses pelabelan, data akan divisualisasikan dalam bentuk grafik untuk memberikan gambaran yang lebih jelas menggunakan *library Matplotlib*. *Matplotlib* adalah pustaka *Python* lintas platform yang dirancang untuk menghasilkan *plot* 2D berkualitas tinggi. Dalam penelitian ini, *library* ini digunakan untuk memvisualisasikan data jaringan, seperti distribusi data normal dan anomali, melalui

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

grafik seperti *histogram* atau grafik batang, sehingga mempermudah analisis dan interpretasi.

4) Tahap Pembagian Data Latih dan Data Uji

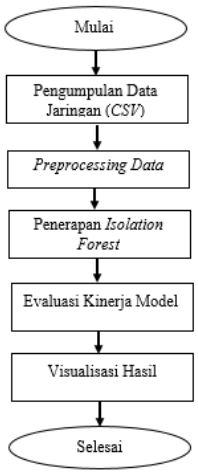
Data *training* adalah data faktual yang telah tersedia sebelumnya, sedangkan data *testing* digunakan untuk mengevaluasi sejauh mana model klasifikasi mampu melakukan klasifikasi dengan benar. Proporsi pembagian antara data *training* dan data *testing* menjadi faktor penting dalam menentukan tingkat akurasi. Kesalahan dalam menetapkan komposisi kedua jenis data ini dapat memengaruhi hasil akurasi dan presisi yang diperoleh [2].

5) Implementasi Isolation Forest

Algoritma *Isolation Forest* digunakan untuk mendeteksi anomali dalam data jaringan dengan cara mengidentifikasi pola distribusi yang berbeda dalam dataset. Model ini bekerja dengan membangun pohon keputusan secara acak untuk memisahkan data, dan mengisolasi data yang dianggap berbeda atau anomali dari data normal. Dalam penelitian ini, algoritma *Isolation Forest* diterapkan pada data jaringan yang telah melalui tahapan *preprocessing*. Evaluasi kinerja model dilakukan untuk mengukur kemampuan algoritma dalam mendeteksi anomali. Alur penerapan *Isolation Forest* dapat dilihat pada gambar berikut ini :

Classification Report:

	precision	recall	f1-score	support
Anomali	0.02	1.00	0.03	22
Normal	1.00	0.80	0.89	6777
accuracy			0.80	6799
macro avg	0.51	0.90	0.46	6799
weighted avg	1.00	0.80	0.89	6799



Gambar 3. Penerapan *Isolation Forest*

III. HASIL DAN PEMBAHASAN

Pengujian yang dilakukan membuktikan bahwa algoritma *Isolation Forest* memiliki kemampuan mendeteksi anomali dengan tingkat akurasi tertinggi mencapai 81% pada pembagian data latih dan uji dengan rasio 80:20. Evaluasi melalui *Confusion Matrix* menunjukkan bahwa model dapat membedakan data normal dan anomali dengan baik. Selain itu, perbedaan rasio pembagian data juga mempengaruhi hasil, di mana akurasi mengalami sedikit penurunan pada rasio 90:10.

A. Hasil

1) Tahap Pengujian Model

Classification Report:

	precision	recall	f1-score	support
Anomali	0.02	1.00	0.03	15
Normal	1.00	0.81	0.89	4518
accuracy			0.81	4533
macro avg	0.51	0.90	0.46	4533
weighted avg	1.00	0.81	0.89	4533

Gambar 4. Hasil Pengujian Model Rasio 70:30

Pada rasio 70:30 akurasi model Isolation Forest mencapai 80%, dengan performa yang sangat baik pada setiap kelas.

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

Precision, recall, dan f1-score untuk semua kelas berkisar antara 0.02 hingga 1.00.

Gambar 5. Hasil Pengujian Model Rasio 80:20

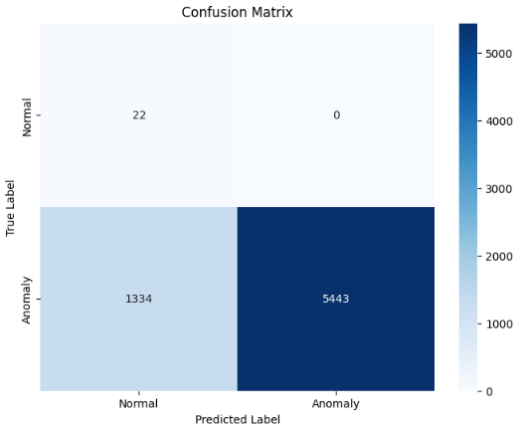
Pada rasio 80:20 akurasi model *Isolation Forest* mencapai 81%, dengan performa yang sangat baik pada setiap kelas. Precision, recall, dan f1-score untuk semua kelas berkisar antara 0.02 hingga 1.00

```
Classification Report:
              precision    recall  f1-score   support

     Anomali       0.01      1.00      0.03         7
      Normal       1.00      0.77      0.87      2260

    accuracy                           0.77      2267
   macro avg       0.51      0.88      0.45      2267
weighted avg       1.00      0.77      0.86      2267
```

Gambar 6. Hasil Pengujian Model Rasio 90:10

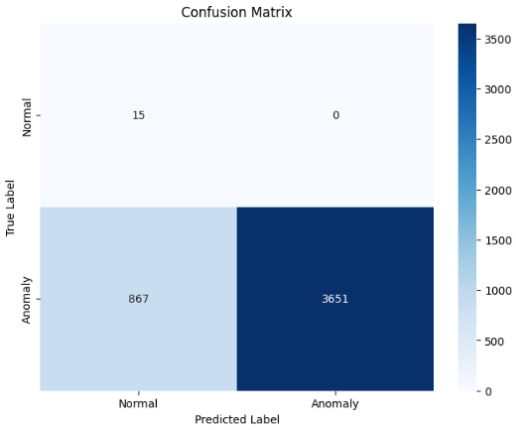
Pada rasio 90:10 akurasi model Isolation Forest mencapai 77%, dengan performa yang sangat baik pada setiap kelas. Precision, recall, dan f1-score untuk semua kelas berkisar antara 0.01 hingga 1.00.

2) Tahap Evaluasi Model



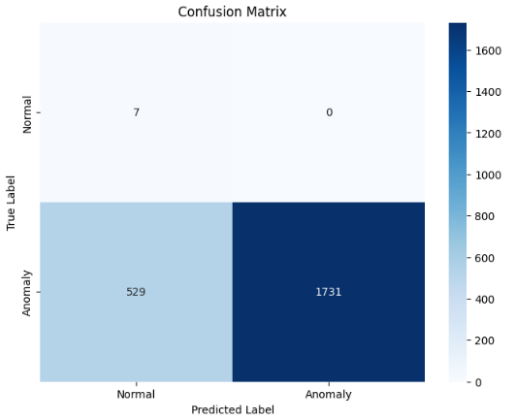
Gambar 7. Hasil Evaluasi Model Rasio 70:30

Confusion Matrix untuk rasio 70:30 menunjukkan bahwa model mampu mengklasifikasikan data dengan tingkat akurasi yang baik. Sebagian besar data berada pada diagonal utama, menunjukkan klasifikasi yang benar untuk setiap kelas (Normal, Anomali).



Gambar 8. Hasil Evaluasi Model Rasio 80:20

Confusion Matrix untuk rasio 80:20 menunjukkan bahwa model mampu mengklasifikasikan data dengan tingkat akurasi yang baik. Sebagian besar data berada pada diagonal utama, menunjukkan klasifikasi yang benar untuk setiap kelas (Normal, Anomali).



Gambar 9. Hasil Evaluasi Model Rasio 90:10

Confusion Matrix untuk rasio 90:10 menunjukkan bahwa model mampu mengklasifikasikan data dengan tingkat akurasi 40 yang baik. Sebagian besar data berada pada diagonal utama, menunjukkan klasifikasi yang benar untuk setiap kelas (Normal, Anomali).

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan tentang Deteksi Anomali dalam Data Jaringan menggunakan Algoritma Isolation Forest, dapat disimpulkan sebagai berikut :

1. Algoritma Isolation Forest menunjukkan kinerja yang sangat baik dalam mendeteksi anomali pada data jaringan. Model ini mampu mengidentifikasi data abnormal secara efisien, bahkan pada dataset berukuran besar, dengan tingkat akurasi hingga 81% ketika menggunkan rasio pembagian data latih dan uji sebesar 80:20.
2. Evaluasi menggunakan Confusion Matrix menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik antara data normal dan anomali.
3. Penelitian ini memberikan kontribusi praktis dengan menerapkan algoritma Isolation Forest untuk mendukung keamanan jaringan. Selain itu, penelitian ini turut memperbanyak refrensi ilmiah dalam bidang deteksi anomali berbasis machine learning.

V. SARAN

Beberapa saran yang dapat diterapkan untuk meningkatkan hasil pada penelitian selanjutnya antara lain :

1. Disarankan untuk peneliti selanjutnya mempertimbangkan penambahan algoritma klasifikasi lain, seperti Decision Tree, Random Forest, dan Naïve Bayes, guna membandingkan kinerja dengan algoritma Isolation Forest yang diterapkan dalam penelitian ini. Hal ini akan memberikan pemahaman yang lebih komprehensif tentang efektivitas setiap algoritma dalam mendeteksi anomali pada data jaringan.
2. Selain itu, penggunaan metode labeling alternatif, seperti pendekatan berbasis statistik atau teknik deteksi anomali lainnya, 42 dapat dipertimbangkan. Dengan mengaplikasikan berbagai pendekatan ini, hasil deteksi anomali dapat lebih tervalidasi dan dibandingkan, yang diharapkan akan meningkatkan akurasi dan kualitas model deteksi anomali yang digunakan.

VI. DAFTAR PUSTAKA

[1] A. Zulfikar, F. A. Rahmani, N. Azizah, D. J. Perbendaharaan, K. Keuangan, and P. Pinang, “Deteksi Anomali Menggunakan Isolation Forest Belanja Barang Persediaan Konsumsi Pada Satuan Kerja Kepolisian Republik Indonesia,” *J. Manaj. Perbendaharaan*, vol. 4, no. 1, pp. 1–15, 2023, doi: 10.33105/jmp.v4i1.435.

[2] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, “Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.

[3] G. Martha, “Pengembangan Algoritma Machine Learning,” vol. 7, pp. 1361–1368, 2024.

[4] S. Situmorang, U. Islam, N. Sumatera, P. Matematika, U. Islam, and N. Sumatera, “Analisis Kinerja Algoritma Machine Learning Dalam Deteksi Anomali Jaringan Sintia Situmorang,” vol. 1, no. 4, 2023.

[5] Milka Wijayanti Sunarto, Dendy Kurniawan, Edy Siswanto, and Haris Ihsanil Huda, “Deteksi Anomali Menggunakan Extended Isolation Forest (Eif),” Tek. J. Ilmu Tek. dan Inform., vol. 1, no. 2, pp. 96–111, 2023, doi: 10.51903/teknik.v1i2.324.

[6] A. Devia and B. Soewito, “Analisis Perbandingan Metode Seleksi Fitur untuk Mendeteksi Anomali pada Dataset CIC-IDS-2018,” J. Teknol. Dan Sist. Inf. Bisnis-JTEKSIS, vol. 5, no. 4, p. 572, 2023, [Online]. Available: http://jurnal.unidha.ac.id/index.php/jteksis

[7] R. M. Imam, P. Sukarno, and M. A. Nugroho, “Deteksi Anomali Jaringan Menggunakan Hybrid Algorithm,”

ESCAF 4th 2025
p-ISSN : 2962-7710
e-ISSN : 3021-8594

2019.
[8] D. Firdaus and R. Rianti, “Dalam Lalu Lintas Jaringan Menggunakan Naive,” vol. 05, no. 02, pp. 140–148, 2023.