Zero-Day Attack Detection Using Autoencoder and XGBoost

Mujibbur Rohman [1)*)], Dharmayanti [2)]

[1)2)]Departemen Teknologi dan Rekayasa, Fakultas Magister Manajemen Sistem Informasi, Universitas Gunadarma
[*)]**Correspondence author**: mujibbur.rohman@gmail.com, Jakarta, Indonesia
DOI: https://doi.org/10.37012/jtik.v12i1.3248

Abstract

Advances in information and communication technology have significantly impacted progress in various sectors, but they have also given rise to increasingly complex network security threats. Cyberattacks such as Distributed Denial of Service (DDoS), ransomware, and software vulnerability exploits continue to increase year after year. Signature-based Intrusion Detection Systems are often ineffective in identifying novel cyberattacks since they rely solely on previously known attack patterns. To address this limitation, this study proposes a hybrid approach that integrates Autoencoders, including Dense and Memory-Augmented variants, with Extreme Gradient Boosting (XGBoost) to enhance zero-day attack detection using the UNSW-NB15 dataset. The research methodology encompasses data exploration, preprocessing with a split-before-transform strategy to prevent information leakage, Autoencoder training to model normal network behavior, reconstruction error computation for anomaly detection under both fixed and adaptive thresholding, and the utilization of these errors as input features for XGBoost classification. Experimental results demonstrate that adaptive thresholding improves F1 performance compared to fixed thresholds, while the hybrid Autoencoder–XGBoost integration achieves a significant performance boost. The proposed model consistently obtained F1 scores above 0.80 and PR-AUC values exceeding 0.81 with a balanced trade-off between precision and recall. These findings confirm that the hybrid approach is more effective, consistent, and adaptive in detecting intrusions, particularly zero-day attacks, and highlight its potential as a robust framework for advancing network security in dynamic threat environments.

Keywords: *Autoencoder, XGBoost, Zero-Day Attacks*

Abstrak

Perkembangan teknologi informasi dan komunikasi telah membawa dampak yang signifikan terhadap kemajuan di berbagai sektor, tetapi di sisi lain juga memunculkan ancaman keamanan jaringan yang semakin kompleks. Serangan siber seperti *Distributed Denial of Service (DDoS)*, ransomware, dan eksploitasi kerentanan perangkat lunak terus meningkat dari tahun ke tahun. *Intrusion Detection System* berbasis tanda tangan terbukti kurang efektif dalam mengenali serangan baru karena keterbatasannya pada pola yang sudah diketahui. Untuk mengatasi masalah ini, penelitian ini mengusulkan pendekatan hibrida yang mengombinasikan Autoencoder, termasuk varian *Dense* dan *Memory-Augmented* Autoencoder, dengan XGBoost guna meningkatkan kemampuan deteksi serangan zero-day pada dataset UNSW-NB15. Metodologi penelitian mencakup tahapan eksplorasi data, *preprocessing* dengan strategi *split-before-transform* untuk mencegah kebocoran informasi, pelatihan Autoencoder dalam memodelkan perilaku normal, perhitungan *reconstruction error* untuk mendeteksi anomali dengan pendekatan *threshold* tetap maupun adaptif, serta pemanfaatan *error* tersebut sebagai fitur masukan bagi XGBoost. Hasil eksperimen menunjukkan bahwa *threshold adaptif* mampu meningkatkan nilai F1 dibanding *threshold* tetap, sementara integrasi Autoencoder dengan XGBoost memberikan peningkatan signifikan pada kinerja deteksi. Model hibrida ini secara konsisten mencapai nilai F1 di atas 0,80 dan PR-AUC lebih dari 0,81 dengan keseimbangan *precision* dan *recall* yang baik. Penelitian ini menegaskan bahwa integrasi Autoencoder dan XGBoost merupakan pendekatan yang efektif, konsisten, dan adaptif untuk menghadapi tantangan deteksi intrusi, khususnya serangan zero-day.

Kata Kunci: Autoencoder, XGBoost, Serangan Zero-Day

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa dampak yang signifikan terhadap kemajuan di berbagai sektor, tetapi di sisi lain juga memunculkan ancaman keamanan jaringan yang semakin kompleks. Serangan siber seperti *Distributed Denial of Service* (DDoS), ransomware, dan eksploitasi kerentanan perangkat lunak terus meningkat dari tahun ke tahun. Salah satu ancaman paling berbahaya adalah serangan zero-day yang memanfaatkan kerentanan yang belum terdokumentasi maupun diperbaiki oleh penyedia sistem, sehingga sulit dideteksi oleh mekanisme pertahanan tradisional. *Intrusion Detection System* (IDS) berbasis tanda tangan terbukti efektif mendeteksi serangan yang sudah dikenal tetapi gagal menghadapi pola serangan baru. Oleh karena itu, pendekatan *anomaly-based detection* dengan *machine learning* menjadi solusi menjanjikan karena mampu mempelajari perilaku normal jaringan dan mengidentifikasi penyimpangan yang mencurigakan.

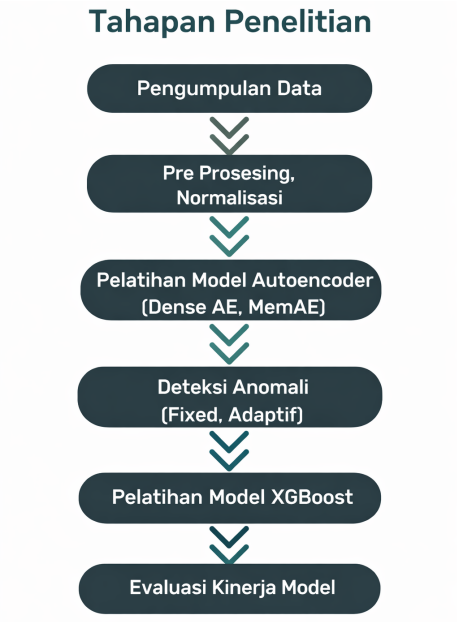
Beberapa penelitian sebelumnya menunjukkan efektivitas *machine learning* dalam meningkatkan performa IDS. Autoencoder sebagai model pembelajaran representasi tanpa label mampu merekonstruksi pola normal lalu lintas jaringan, sehingga penyimpangan rekonstruksi dapat dimanfaatkan sebagai sinyal anomali. Pengembangan *Memory-Augmented* Autoencoder (MemAE) terbukti meningkatkan akurasi deteksi anomali dengan memanfaatkan modul memori untuk membedakan data normal dan abnormal secara lebih tegas (Gong *et al*. 2019). Penggunaan algoritma *boosting* seperti XGBoost dikenal unggul dalam menangani data tidak seimbang serta mampu meningkatkan *precision* dan *recall* dalam skenario deteksi intrusi. Penelitian terdahulu (Dai *et al*. 2024) mengombinasikan autoencoder dengan XGBoost dan menunjukkan peningkatan performa pada data *unseen*, namun masih terbatas pada arsitektur autoencoder standar tanpa memanfaatkan varian berbasis memori.

Berdasarkan latar belakang dan hasil penelitian terdahulu tersebut, penelitian ini dilakukan untuk mengeksplorasi efektivitas integrasi *Dense* Autoencoder dan *Memory-Augmented* Autoencoder dengan XGBoost dalam mendeteksi serangan zero-day menggunakan dataset UNSW-NB15. Tujuan utama penelitian ini adalah menganalisis keterbatasan IDS berbasis tanda tangan, mengevaluasi metode *threshold* tetap dan adaptif,

serta membuktikan bahwa pendekatan hibrida mampu meningkatkan akurasi, *precision*, *recall*, dan F1-*score* dalam mendeteksi intrusi, khususnya serangan zero-day.

METODE PENELITIAN

Penelitian ini difokuskan pada pengembangan sistem deteksi intrusi berbasis *machine learning* untuk menghadapi serangan zero-day, *dataset* yang digunakan adalah UNSW-NB15 dengan cakupan lebih dari 250 ribu baris data lalu lintas jaringan. Ruang lingkup dibatasi pada eksperimen berbasis simulasi menggunakan Python dan Jupyter Notebook, tanpa implementasi langsung pada infrastruktur IDS nyata. Model yang diteliti meliputi *Dense* Autoencoder (AE), *Memory-Augmented* Autoencoder (MemAE), diintegrasikan dengan XGBoost sebagai pendekatan hibrida.



Gambar 1. Tahapan Penelitian

Metode penelitian dilakukan melalui 6 tahap seperti Gambar 1, diantaranya:

1. **Pengumpulan Data**

   Tahap awal penelitian dilakukan dengan memanfaatkan dataset UNSW-NB15 yang berisi lalu lintas jaringan normal dan anomali. Data diperiksa melalui audit log, statistik inti, distribusi fitur, serta analisis *missing values* dan duplikasi.

2. **Preprocessing dan Normalisasi**

   Dataset dibagi menjadi *train*, *validation*, dan *test*, dengan strategi *split-before-transform* untuk mencegah data *leakage*. Fitur yang berpotensi membocorkan label, seperti id dan attack_cat, dieliminasi untuk menjaga validitas eksperimen. Imputasi median diterapkan untuk fitur numerik dan token khusus untuk kategorikal. Selanjutnya dilakukan *encoding* hemat memori dan normalisasi dengan StandardScaler agar distribusi fitur seimbang dan stabil pada seluruh subset data.

3. **Pelatihan Autoencoder (*Dense* AE dan MemAE)**

   Model *Dense* AE dilatih menggunakan arsitektur *encoder–decoder* berbasis lapisan penuh, sementara MemAE menambahkan modul memori untuk menyimpan representasi pola normal. Kedua model dilatih dengan fungsi *loss* MSE. Teknik *Early Stopping* digunakan untuk menghentikan pelatihan pada titik optimal, sedangkan Model *Checkpoint* menyimpan bobot terbaik.

4. **Deteksi Anomali (*Threshold* Tetap dan Adaptif)**

   Reconstruction error yang dihasilkan Autoencoder dipetakan menjadi skor anomali. Dua metode *threshold* digunakan yaitu *threshold* tetap berdasarkan persentil distribusi *error* data normal, serta *threshold* adaptif berdasarkan optimasi F1-*score* pada *validation* set. Pendekatan ini memungkinkan evaluasi keseimbangan antara *False Positive Rate* dan *recall*.

5. **Integrasi dengan XGBoost**

   Reconstruction error dari AE/MemAE dijadikan fitur utama bagi XGBoost, dengan kemungkinan penambahan subset fitur numerik. XGBoost memanfaatkan *gradient boosting decision trees* untuk mempelajari pola non-linear, sehingga meningkatkan akurasi klasifikasi intrusi, termasuk pada data serangan zero-day. Analisis *threshold-vs-metrics* digunakan untuk memahami pengaruh variasi ambang terhadap performa model. Hasilnya dibandingkan antara AE murni, *thresholding*, dan integrasi AE–XGBoost untuk menentukan pendekatan paling efektif.

6. **Evaluasi Kinerja Model**

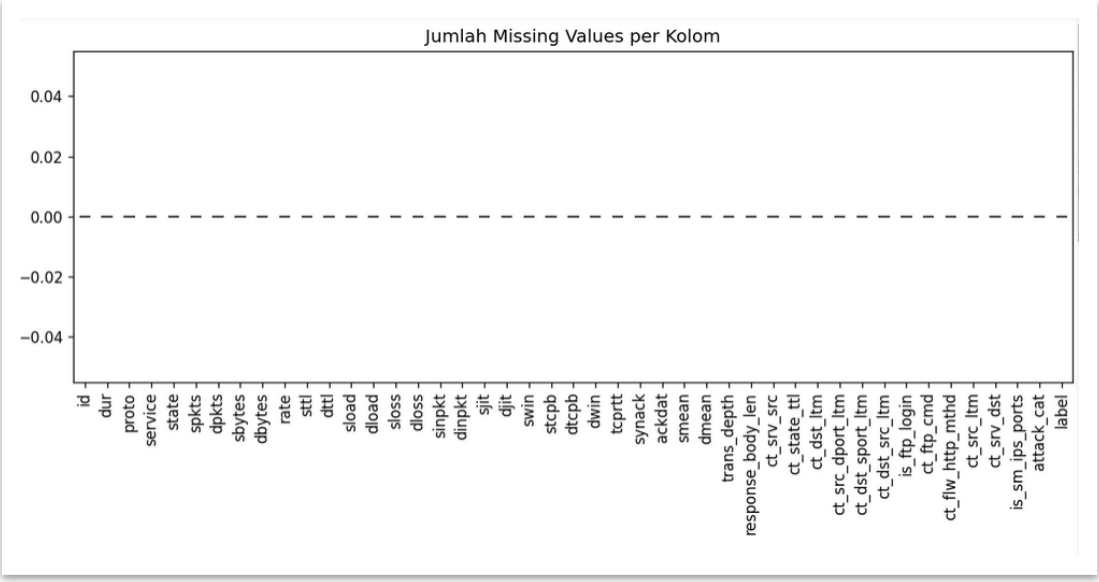
   Model dievaluasi pada data uji menggunakan visualisasi: ROC-AUC, dan PR-AUC.

HASIL DAN PEMBAHASAN

```
Loading: UNSW_NB15_training-set.csv
Loading: UNSW_NB15_testing-set.csv
Sukses memuat dataset: shape=(257673, 45)
```

Gambar 2. Memuat Data

Tahap pengumpulan data menggunakan dataset UNSW-NB15 berhasil dimuat dengan 257.673 record dan 45 atribut seperti terlihat di Gambar 2.



Gambar 3. *Bar Chart Missing Values*

Tahap *preprocessing* dan normalisasi dilakukan pemeriksaan kualitas data, hasilnya menunjukkan tidak terdapat *missing values* seperti terlihat di Gambar 3.

```
Distribusi Label (Persentase):
label
1    63.907744
0    36.092256
Name: proportion, dtype: float64
```

Gambar 4. Distribusi Serangan vs Normal

Distribusi label memperlihatkan ketidakseimbangan kelas dengan 63,9% data berupa serangan (*attack*) dan 36,1% berupa trafik normal terlihat di Gambar 4.

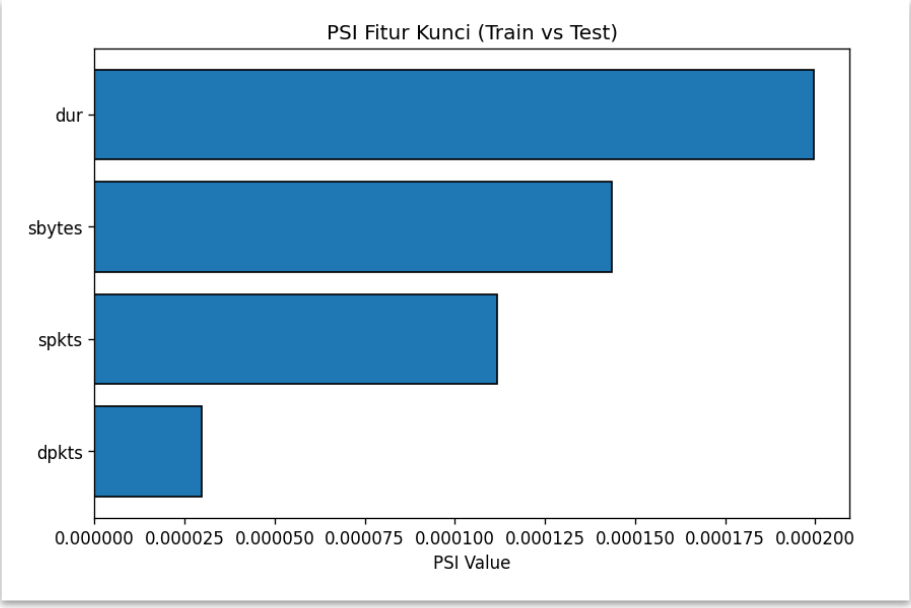
```
Jumlah kolom sebelum eliminasi: 45
Daftar kolom sebelum eliminasi: ['id', 'dur', 'pr
'sloss', 'dloss', 'sinpkt', 'dinpkt', 'sjit', 'dj
ponse_body_len', 'ct_srv_src', 'ct_state_ttl', 'c
w_http_mthd', 'ct_src_ltm', 'ct_srv_dst', 'is_sm_

Jumlah kolom sesudah eliminasi: 43
Daftar kolom sesudah eliminasi: ['dur', 'proto',
', 'dloss', 'sinpkt', 'dinpkt', 'sjit', 'djit', '
body_len', 'ct_srv_src', 'ct_state_ttl', 'ct_dst_
_mthd', 'ct_src_ltm', 'ct_srv_dst', 'is_sm_ips_po

Kolom yang dihapus: ['id', 'attack_cat']
```

Gambar 5. Eliminasi Kolom Beresiko

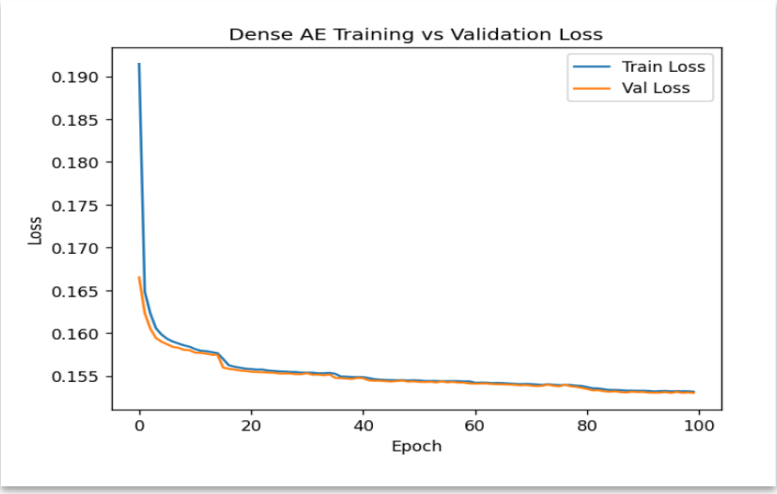
Proses prapemrosesan mencakup eliminasi kolom berisiko (id, attack_cat), *encoding* kategorikal (proto, service, state), serta normalisasi numerik terlihat di Gambar 5.



Gambar 6. Population *Stability Index* (PSI)

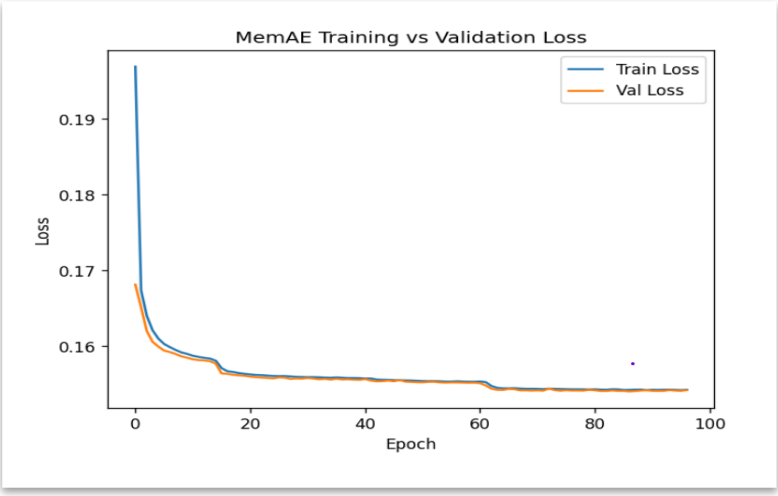
Gambar 6 Analisis *Population Stability Index* (PSI) menunjukkan seluruh fitur inti (dur, spkts, dpkts, sbytes) memiliki nilai < 0,10, sehingga distribusi data stabil dan layak digunakan untuk pelatihan model.

Tahap pelatihan Autoencoder (*Dense* AE dan MemAE), model *Dense* Autoencoder (AE) dan *Memory*-Augmented Autoencoder (MemAE) dilatih dengan mekanisme *early stopping*.



Gambar 7. Dense AE *Validation Loss*

Pada *Dense* AE di Gambar 7, *train loss* berada di sekitar 0,190 dan kemudian turun secara bertahap hingga mencapai sekitar 0,155 pada epoch ke-100. Pola *Validation loss* mengikuti pola Dense AE, dimulai sekitar 0,165 dan menurun hingga sekitar 0,155. Pola ini menunjukkan bahwa *Dense* AE berhasil mencapai konvergensi stabil, dengan gap yang sangat kecil antara *train loss* dan *validation loss*, sehingga risiko *overfitting* rendah.



Gambar 8. MemAE *Validation Loss*

Pada MemAE di Gambar 8, *train loss* awal sekitar 0,195 dan menurun hingga mendekati 0,157, sedangkan validation *loss* dimulai pada 0,167 dan stabil di sekitar 0,158 pada akhir epoch ke-100. Kurva keduanya menunjukkan tren yang berdekatan dan relatif konvergen. Hal ini mengindikasikan bahwa MemAE mampu merepresentasikan pola normal dengan baik, bahkan lebih stabil karena adanya modul memori yang membantu rekonstruksi data normal secara lebih akurat.

Tabel 1. Autoencoder dengan *Threshold* Tetap 70 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (70)	0.561439	0.217623	0.313664	0.345062
MEMAE - FIXED (70)	0.552489	0.20987	0.30419	0.338718
DENSE AE — ADAPTIVE	0.639069	0.99996	0.779782	0.345062
MEMAE — ADAPTIVE	0.639064	0.999939	0.779773	0.338718

Tabel 2. Autoencoder dengan *Threshold* Tetap 75 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (75)	0.560773	0.182017	0.274829	0.34609
MEMAE - FIXED (75)	0.559071	0.181329	0.27384	0.341047
DENSE AE — ADAPTIVE	0.639069	0.99996	0.779782	0.34609
MEMAE — ADAPTIVE	0.639069	0.99996	0.779782	0.341047

Tabel 3. Autoencoder dengan *Threshold* Tetap 80 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (80)	0.530377	0.128294	0.206611	0.339836
MEMAE - FIXED (80)	0.545476	0.138395	0.220776	0.337646
DENSE AE — ADAPTIVE	0.639077	1.0	0.779989	0.339836
MEMAE — ADAPTIVE	0.639078	1.0	0.779802	0.337646

Tahap deteksi anomali dilakukan dengan *threshold* tetap (persentil 70, 75, dan 80) serta *threshold* adaptif (optimasi F1 pada *validation* set) seperti pada Tabel 1, Tabel 2, Tabel 3. Pada *threshold* tetap (persentil 70–80) *Dense* AE menghasilkan *precision* antara 0,53–0,56, *recall* hanya sekitar 0,12–0,22, dan F1 di kisaran 0,20–0,31. Pada MemAE menunjukkan hasil yang hampir sama dengan *precision* 0,55, *recall* 0,13–0,20, dan F1 0,22–0,30. Performa *threshold* tetap buruk karena *recall* sangat rendah, menandakan banyak serangan tidak terdeteksi. Pada *threshold* adaptif, baik *Dense* AE maupun MemAE konsisten mencapai *precision* sekitar 0,639, *recall* mendekati 1,0, dan F1 sekitar 0,7798. Hal ini

menunjukkan bahwa *threshold* adaptif jauh lebih efektif dalam menyeimbangkan precision dan *recall*, serta secara signifikan meningkatkan deteksi serangan zero-day dibandingkan *threshold* tetap.

Tabel 4. Autoencoder+*XGBoost* dengan *Threshold* Tetap 70 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (70)	0.561439	0.217623	0.313664	0.345062
MEMAE - FIXED (70)	0.552489	0.20987	0.30419	0.338718
DENSE AE — ADAPTIVE	0.639069	0.99996	0.779782	0.345062
MEMAE — ADAPTIVE	0.639064	0.999939	0.779773	0.338718
XGB — AE-FIXED [TEST]	0.721914	0.912089	0.805935	0.752838
XGB — AE-ADAPTIF [TEST]	0.704602	0.953099	0.810225	0.752838
XGB — MEMAE-FIXED [TEST]	0.716917	0.925671	0.808029	0.739763
XGB — MEMAE-ADAPTIF [TEST]	0.709552	0.947026	0.811268	0.739763

Tabel 5. Autoencoder+*XGBoost* dengan *Threshold* Tetap 75 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (75)	0.560773	0.182017	0.274829	0.34609
MEMAE - FIXED (75)	0.559071	0.181329	0.27384	0.341047
DENSE AE — ADAPTIVE	0.639069	0.99996	0.779782	0.34609
MEMAE — ADAPTIVE	0.639069	0.99996	0.779782	0.341047
XGB — AE-FIXED [TEST]	0.719144	0.924193	0.808876	0.757362
XGB — AE-ADAPTIF [TEST]	0.708262	0.953362	0.812735	0.757362
XGB — MEMAE-FIXED [TEST]	0.7198	0.924861	0.809547	0.737434
XGB — MEMAE-ADAPTIF [TEST]	0.702053	0.965852	0.813091	0.737434

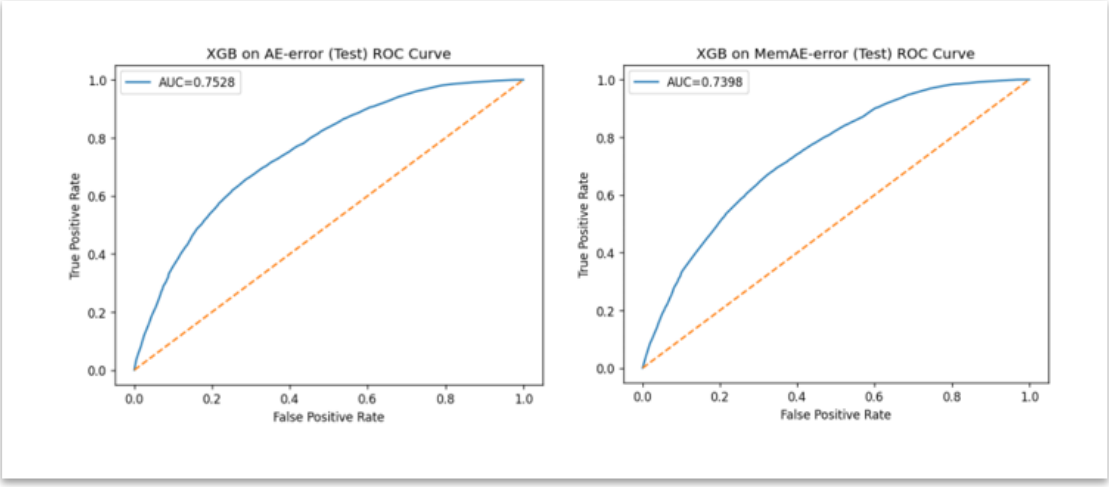
Tabel 6. Autoencoder+*XGBoost* dengan *Threshold* Tetap 80 dan *Threshold* Adaptif

MODEL	PRECISION	RECALL	F1	ROC-AUC
DENSE AE - FIXED (80)	0.530377	0.128294	0.206611	0.339836
MEMAE - FIXED (80)	0.545476	0.138395	0.220776	0.337646
DENSE AE — ADAPTIVE	0.639077	1.0	0.779989	0.339836
MEMAE — ADAPTIVE	0.639078	1.0	0.779802	0.337646
XGB — AE-FIXED [TEST]	0.716366	0.926076	0.807833	0.752197
XGB — AE-ADAPTIF [TEST]	0.69901	0.963605	0.810253	0.752197
XGB — MEMAE-FIXED [TEST]	0.714091	0.934294	0.809485	0.736888
XGB — MEMAE-ADAPTIF [TEST]	0.709653	0.945812	0.810888	0.736888

Tahap integrasi dengan XGBoost menunjukkan peningkatan signifikan dibandingkan *threshold* murni, terlihap pada Tabel 4, Tabel 5, Tabel 6. Pada *Dense* AE +

XGBoost *Precision* = 0.708262, *Recall* = 0.953362, F1 = 0.812735, ROC-AUC = 0.757362. Pada MemAE + XGBoost *Precision* = 0.702053, *Recall* = 0.965852, F1 = 0.813091, ROC-AUC = 0.737434. Kedua pendekatan konsisten menghasilkan F1 di atas 0,81, menandakan keseimbangan yang baik antara *precision* dan *recall* walaupun ROC-AUC tidak terlalu tinggi (sekitar 0,74–0,75). Hasil ini menegaskan bahwa integrasi AE/MemAE dengan XGBoost lebih efektif dibanding *threshold* murni dalam mendeteksi serangan zero-day, khususnya pada kondisi data yang tidak seimbang.

Tahap evaluasi kinerja model melengkapi tahap integrasi XGBoost dengan gambar visual pendukung meliputi kurva ROC, kurva PR-AUC



Gambar 9. Kurva ROC pada AE dan MemAE

Kurva PR Gambar 9 menjelaskan seberapa baik model mempertahankan keseimbangan antara *Precision* dan *Recall* pada kondisi data tidak seimbang, dengan hasil yang menunjukkan performa tinggi (AUC > 0.81).

Pembahasan

Hasil penelitian *menunjukkan* bahwa pendekatan hibrida Autoencoder dengan XGBoost mampu meningkatkan efektivitas deteksi intrusi, khususnya dalam menghadapi serangan zero-day. Integrasi reconstruction *error* dari Autoencoder ke dalam XGBoost secara konsisten menghasilkan nilai F1 di atas 0.80 dan PR-AUC lebih dari 0.81, menandakan keseimbangan yang baik antara *precision* dan *recall* meskipun dataset tidak seimbang.

Penggunaan *threshold* adaptif terbukti lebih unggul dibandingkan *threshold* tetap. *Threshold* tetap memberikan kontrol False *Positive Rate* (FPR) yang stabil tetapi tidak mampu mengoptimalkan F1-*score*. Sebaliknya, *threshold* adaptif yang dikalibrasi pada data validasi mampu menyeimbangkan *precision* dan *recall* secara lebih efektif. Hal ini sejalan dengan penelitian sebelumnya (Zoppi *et al*. 2021) yang menunjukkan keunggulan *threshold* adaptif dalam menghadapi distribusi *reconstruction error* yang saling tumpang tindih.

Perbandingan kinerja antara *Dense* AE dan MemAE menunjukkan bahwa MemAE lebih *robust* dalam mendeteksi anomali. Keunggulan MemAE disebabkan oleh adanya modul memori yang memungkinkan penyimpanan representasi pola normal yang lebih kaya. Temuan ini mendukung penelitian Gong *et al*. (2019) yang menegaskan bahwa MemAE lebih efektif membedakan data normal dan abnormal dibanding *Dense* AE.

Nilai ROC-AUC model hibrida berada di kisaran 0.74–0.75, tapi performa PR-AUC yang melebihi 0.81 lebih relevan untuk kasus IDS dengan data tidak seimbang. PR-AUC menekankan performa pada kelas minoritas (serangan) yang menjadi fokus utama dalam sistem deteksi intrusi. Temuan ini selaras dengan literatur mutakhir (Dai *et al*. 2024) yang menekankan pentingnya PR-AUC sebagai metrik utama dalam evaluasi IDS berbasis *machine learning*.

Secara keseluruhan, hasil penelitian ini memperkuat pemahaman bahwa integrasi metode *unsupervised* (Autoencoder) dengan *supervised boosting classifier* (XGBoost) dapat menghasilkan sistem deteksi intrusi yang lebih adaptif, konsisten, dan efektif dalam mendeteksi serangan zero-day. Model hibrida ini memiliki potensi besar untuk diimplementasikan dalam sistem IDS modern, dengan tetap memperhatikan efisiensi komputasi dan skalabilitas pada trafik jaringan yang besar.

KESIMPULAN DAN REKOMENDASI

Penelitian ini membuktikan efektivitas pendekatan hibrida antara Autoencoder (Dense AE dan Memory-Augmented AE) dengan XGBoost dalam meningkatkan kemampuan sistem deteksi intrusi, terutama terhadap serangan zero-day pada dataset UNSW-NB15. Melalui serangkaian eksperimen terstruktur mulai dari pra-pemrosesan bebas kebocoran (split-before-transform), pelatihan model Autoencoder untuk mempelajari

perilaku normal jaringan, penentuan ambang anomali menggunakan metode tetap dan adaptif, hingga klasifikasi lanjutan menggunakan XGBoost, penelitian ini berhasil menunjukkan peningkatan signifikan dalam performa deteksi anomali yang kompleks dan belum dikenal sebelumnya.

Hasil eksperimen menunjukkan bahwa penggunaan *threshold* adaptif secara konsisten mengungguli *threshold* tetap, dengan peningkatan F1-score yang signifikan karena mampu menyeimbangkan *precision* dan *recall* secara dinamis. Hal ini menunjukkan bahwa mekanisme adaptif lebih responsif terhadap variasi distribusi data dan mampu meminimalkan tingkat *false negative* pada serangan yang belum pernah muncul sebelumnya. Selain itu, Memory-Augmented Autoencoder (MemAE) terbukti lebih unggul dibandingkan Dense AE karena kemampuannya menyimpan dan membedakan representasi pola normal secara lebih efisien melalui modul memori, sehingga menghasilkan rekonstruksi yang lebih akurat dan stabil.

Integrasi *reconstruction error* dari Autoencoder ke dalam XGBoost memberikan peningkatan performa yang signifikan, di mana model hibrida AE–XGBoost mencapai F1-score di atas 0,80 dan PR-AUC lebih dari 0,81, dengan tingkat keseimbangan yang optimal antara *precision* dan *recall*. Meskipun nilai ROC-AUC berada pada kisaran 0,74–0,75, hal ini masih dianggap memadai karena PR-AUC merupakan metrik yang lebih representatif dalam konteks dataset tidak seimbang seperti IDS. Dengan demikian, model ini tidak hanya efektif dalam mengenali anomali yang sudah dikenal, tetapi juga adaptif terhadap serangan baru tanpa memerlukan pembaruan tanda tangan (*signature*).

Secara konseptual dan empiris penelitian ini menegaskan bahwa kombinasi *unsupervised learning* (Autoencoder) dan *supervised boosting* (XGBoost) mampu menghasilkan sistem *Intrusion Detection System* (IDS) yang lebih tangguh, adaptif, dan *generalizable* terhadap variasi serangan zero-day. Pendekatan ini tidak hanya memperluas cakupan deteksi tetapi juga meningkatkan efisiensi proses pembelajaran melalui pemanfaatan rekonstruksi error sebagai fitur semantik yang bermakna bagi model klasifikasi.

Selain kontribusi empiris penelitian ini juga memberikan dasar teoritis yang kuat bagi pengembangan IDS modern berbasis kecerdasan buatan (AI-driven IDS) yang mampu

beradaptasi dengan dinamika ancaman siber. Dengan performa tinggi dan stabilitas model yang baik, sistem hibrida Autoencoder–XGBoost berpotensi untuk diimplementasikan pada lingkungan produksi *real-time* sebagai komponen utama dalam ekosistem *cyber threat intelligence* dan *adaptive anomaly detection*.

Penelitian *lebih* lanjut integrasi model ini dapat diperluas dengan penambahan arsitektur *deep learning* yang lebih kompleks seperti Variational Autoencoder (VAE), Transformer, serta penerapan Explainable AI (XAI) untuk meningkatkan transparansi keputusan sistem. Implementasi pada trafik *real-time* dan *multi-dataset benchmark* juga akan menjadi langkah penting untuk menilai skalabilitas, efisiensi komputasi, serta kemampuan generalisasi lintas domain. Dengan demikian, penelitian ini tidak hanya memberikan solusi konkret terhadap tantangan deteksi serangan zero-day, tetapi juga membuka arah baru bagi pengembangan sistem keamanan siber otonom dan adaptif berbasis pembelajaran mesin.

REFERENSI

Ahmad, R, Alsmadi, I, Alhamdani, W, & ... (2023). Zero-day attack detection: a systematic literature review. *Artificial Intelligence* ..., Springer, https://doi.org/10.1007/s10462-023-10437-z

Alshehri, A, Badr, MM, Baza, M, & Alshahrani, H (2024). Deep anomaly detection framework utilizing federated learning for electricity theft zero-day cyberattacks. *Sensors*, mdpi.com, https://www.mdpi.com/1424-8220/24/10/3236

Alzaylaee, MK (2025). Enhancing Cybersecurity Through Artificial Intelligence: A Novel Approach to Intrusion Detection.. ... *Journal of Advanced Computer Science & ...*, search.ebscohost.com,

Awadia, OA El, & Salem, SA (2023). Unveiling the Unseen: Leveraging Zero-Day Attack Detection Using Unsupervised and Semi-Supervised Learning. *2023 33rd International Conference ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/10969349/

Aarthi, C, Saranya, K, Saranya, N, & Ponlatha, S (2024). Secured cyber-internet security in intrusion detection with machine learning techniques. *Int. J. Comput. Exp. Sci. Eng*

Babaey, V, & Faragardi, HR (2025). Detecting Zero-Day Web Attacks with an Ensemble of LSTM, GRU, and Stacked Autoencoders. *Computers*, mdpi.com, https://www.mdpi.com/2073-431X/14/6/205

Barnard, P, Marchetti, N, & DaSilva, LA (2022). Robust network intrusion detection through explainable artificial intelligence (XAI). *IEEE Networking Letters*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/9807332/

Binbusayyis, A, & Vaiyapuri, T (2021). Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM. *Applied Intelligence*, Springer, https://doi.org/10.1007/S10489-021-02205-9

Dai, Z, Por, LY, Chen, YL, Yang, J, Ku, CS, & ... (2024). An intrusion detection model to detect zero-day attacks in unseen data using machine learning. *PloS one*, journals.plos.org, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0308469

DAS, A, & PRAMOD, S (2022). An Enhanced Optimization Model With Ensemble Autoencoder For Zero-Day Attack Detection. *Journal of Theoretical and Applied Information ...*, jatit.org, https://www.jatit.org/volumes/Vol100No22/26Vol100No22.pdf

Diallo, AF, & Patras, P (2023). Deciphering clusters with a deterministic measure of clustering tendency. *IEEE Transactions on Knowledge and ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/10227568/

Dini, P, Elhanashi, A, Begni, A, Saponara, S, Zheng, Q, & ... (2023). Overview on intrusion detection systems design exploiting machine learning for networking cybersecurity. *Applied Sciences*, mdpi.com, https://www.mdpi.com/2076-3417/13/13/7507

Gong, D, Liu, L, Le, V, Saha, B, & ... (2019). Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. *Proceedings of the ...*, openaccess.thecvf.com,

Husseini, F El, Noura, H, Salman, O, & ... (2024). Advanced Machine Learning Approaches for Zero-Day Attack Detection: A Review. *2024 8th Cyber ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/10851751/

Hidayat, I, Ali, MZ, & Arshad, A (2023). Machine learning-based intrusion detection system: an experimental comparison. *Journal of Computational and ...*, ojs.bonviewpress.com, http://ojs.bonviewpress.com/index.php/JCCE/article/view/270

Hindy, H, Atkinson, R, Tachtatzis, C, Colin, JN, Bayne, E, & ... (2020). Utilising deep learning techniques for effective zero-day attack detection. *Electronics*, mdpi.com, https://www.mdpi.com/2079-9292/9/10/1684

Id-SIRTII. (2025). "LANSKAP KEAMANAN SIBER INDONESIA."

Islam, M, Dukyil, AS, Alyahya, S, & Habib, S (2023). An IoT enable anomaly detection system for smart city surveillance. *Sensors*, mdpi.com, https://www.mdpi.com/1424-8220/23/4/2358

Lu, K (2024). Network anomaly traffic analysis. *Academic Journal of Science and Technology*

Minghui, G, Hang, Z, Li, M, Zhijun, Z, Kai, L, & ... (2022). Research on intrusion detection model based on DAE-XGBoost. *2022 IEEE 10th ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/10006598/

Mohamed, AA, Al-Saleh, A, Sharma, SK, & Tejani, GG (2025). Zero-day exploits detection with adaptive WavePCA-Autoencoder (AWPA) adaptive hybrid exploit detection network (AHEDNet). *Scientific Reports*, nature.com, https://www.nature.com/articles/s41598-025-87615-2

More, S, Idrissi, M, Mahmoud, H, & Asyhari, AT (2024). Enhanced intrusion detection systems performance with UNSW-NB15 data analysis. *Algorithms*, mdpi.com, https://www.mdpi.com/1999-4893/17/2/64

Muneer, S, Farooq, U, Athar, A, & ... (2024). A critical review of artificial intelligence based approaches in intrusion detection: A comprehensive analysis. *Journal of ...*, Wiley Online Library, https://doi.org/10.1155/2024/3909173

Nassif, AB, Talib, MA, Nasir, Q, & Dakalbab, FM (2021). Machine learning for anomaly detection: A systematic review. *Ieee Access*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/9439459/

Oettl, FC, Oeding, JF, Feldt, R, Ley, C, & ... (2024). The artificial intelligence advantage: Supercharging exploratory data analysis. *Knee Surgery ...*, Wiley Online Library, https://doi.org/10.1002/ksa.12389

Neto, MP, & Paulovich, FV (2022). Multivariate data explanation by jumping emerging patterns visualization. *IEEE Transactions on Visualization ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/9956758/

Roshan, K, & Zafar, A (2021). An optimized auto-encoder based approach for detecting zero-day cyber-attacks in computer network. *2021 5th International Conference on ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/9702437/

Sarhan, M, Layeghy, S, Gallagher, M, & ... (2023). From zero-shot machine learning to zero-day attack detection. *International Journal of ...*, Springer, https://doi.org/10.1007/s10207-023-00676-0

Solunke, P, Guardieiro, V, Rulff, J, & ... (2024). Mountaineer: Topology-Driven Visual Analytics for Comparing Local Explanations. *... on Visualization and ...*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/10570334/

Vaiyapuri, T, & Binbusayyis, A (2020). Application of deep autoencoder as an one-class classifier for unsupervised network intrusion detection: a comparative evaluation. *PeerJ Computer Science*, peerj.com, https://peerj.com/articles/cs-327/

Yazdizadeh, T, Hassani, S, & Branco, P (2022). Intrusion detection using ensemble models. *Joint European Conference on ...*, Springer, https://doi.org/10.1007/978-3-031-23633-4_11

Zahoora, U, Rajarajan, M, Pan, Z, & Khan, A (2022). Zero-day ransomware attack detection using deep contractive autoencoder and voting based ensemble classifier. *Applied Intelligence*, Springer, https://doi.org/10.1007/s10489-022-03244-6

Zhou, KQ (2022). Zero-day vulnerabilities: Unveiling the threat landscape in network security. *Mesopotamian Journal of CyberSecurity*, journals.mesopotamian.press, https://journals.mesopotamian.press/index.php/CyberSecurity/article/view/97

Zoppi, T, Ceccarelli, A, & Bondavalli, A (2021). Unsupervised algorithms to detect zero-day attacks: Strategy and application. *Ieee Access*, ieeexplore.ieee.org, https://ieeexplore.ieee.org/abstract/document/9461213/