

CatBoost untuk Deteksi Worm dengan Dataset UNSW-NB15 menggunakan MDI-VAE Untuk Memecahkan Masalah Label yang Tidak Seimbang

Sultan
Fakultas Informatika
Universitas Telkom
Bandung
sultantann@students.telkomuniversity.ac.id

Aji Gautama Putrada
Fakultas Informatika
Universitas Telkom
Bandung
ajigps@telkomuniversity.ac.id

Ikke Dian Oktaviani
Fakultas Informatika
Universitas Telkom
Bandung
oktavianiid@telkomuniversity.ac.id

Abstrak

Apa. Label serangan worm ada pada dataset UNSW-NB15, tetapi kondisi ketidakseimbangannya parah. Tujuan penelitian ini adalah untuk menganalisis efektivitas variational autoencoder (VAE) dalam mengatasi kondisi ketidakseimbangan di dataset UNSW-NB15.

Mengapa. Studi ini menggunakan salah satu kelas minoritas dalam dataset UNSW-NB15, yaitu kelas worm. Pada tahap selanjutnya, studi ini melakukan seleksi fitur menggunakan mean decrease in impurity (MDI) dengan pengklasifikasi ExtraTrees (ETC) sebagai dasar dan membandingkannya dengan metode terkini, random forest (RF).

Bagaimana. Kami menggunakan akurasi, presisi, recall, dan F1-score untuk membandingkan kedua metode seleksi fitur tersebut. Kemudian, kami mengusulkan model ensemble berdasarkan category boosting (CatBoost) dan membandingkan performanya dengan adaptive boosting (AdaBoost) menggunakan sensitivitas, spesifisitas, dan geometric mean (G-mean). Pada tahap akhir, kami mengevaluasi efek VAE dalam menghasilkan data sintetis untuk menyeimbangkan data pelatihan.

Hasil utama. Hasil pengujian menunjukkan bahwa CatBoost yang menggunakan MDI ETC mengungguli MDI RF, mencapai skor presisi dan G-mean masing-masing sebesar 0,80 dan 0,86, dibandingkan dengan 0,61 dan 0,68. CatBoost mengungguli AdaBoost dalam hal sensitivitas dan G-mean, mencapai 0,97 dan 0,99, dibandingkan dengan 0,67 dan 0,82 untuk AdaBoost. Terakhir, penerapan VAE tidak memiliki efek signifikan pada AUC CatBoost, sedangkan MDI+VAE meningkatkan skor tersebut dari 0,987 menjadi 0,995.

Kata kunci : sistem deteksi intrusi, ketidakseimbangan kelas, variational autoencoder, UNSW NB15, peningkatan boosting.

Abstract

Worm attack labels are present in the UNSW-NB15 dataset, but the imbalance condition is severe. The purpose of this study is to analyze the effectiveness of the variational autoencoder (VAE) in addressing the imbalance condition in the UNSW-NB15 dataset. This study uses one of the minority labels in the UNSW-NB15 dataset, the worm label. In the next stage, this study performed feature selection using mean decrease in impurity (MDI) with a baseline ExtraTrees classifier (ETC) and compared it with the state-of-the-art method, random forest (RF). We used accuracy, precision, recall, and F1-score to compare the two feature selection methods. We then proposed an ensemble model based on category boosting (CatBoost) and compared its performance with adaptive boosting (AdaBoost) using sensitivity, specificity, and the geometric mean (G-mean). In the final stage, we evaluated the effect of VAE in generating synthetic data to balance the training data. Test

results showed that CatBoost using MDI ETC outperformed MDI RF, achieving precision and G-mean scores of 0.80 and 0.86, respectively, compared with 0.61 and 0.68, respectively. CatBoost outperformed AdaBoost in sensitivity and G-mean, achieving 0.97 and 0.99, respectively, compared with 0.67 and 0.82 for AdaBoost. Finally, the application of VAE had no significant effect on CatBoost's AUC, whereas MDI+VAE increased that score from 0.987 to 0.995.

Keywords: intrusion detection systems, class imbalance, variational autoencoder, UNSW-NB15, categorical boosting

1. Pendahuluan

Latar Belakang

Kemajuan teknologi saat ini telah menyebabkan banyak serangan siber terhadap infrastruktur jaringan [1]. Sistem Deteksi Intrusi (IDS) merupakan komponen penting dari arsitektur keamanan jaringan, yang mampu mendeteksi aktivitas jaringan anomali secara real-time. Namun ketidakseimbangan kelas yang dihasilkan menimbulkan tantangan tersendiri bagi sistem IDS berbasis deep learning [2]. Model yang dilatih pada dataset yang tidak seimbang cenderung menunjukkan bias terhadap kelas mayoritas, hal ini dapat mengakibatkan tingkat deteksi yang rendah untuk beberapa jenis serangan langka.

Beberapa penelitian telah mengembangkan sistem deteksi intrusi (IDS) cerdas menggunakan dataset UNSW-NB15. Xie et al. [3] Perlu dicatat bahwa dataset UNSW-NB15 ini mewakili lalu lintas jaringan modern dengan pola serangan realistis dan mencakup sembilan kategori serangan ditambah lalu lintas normal. Namun, dataset ini menunjukkan ketidakseimbangan kelas yang ekstrem, dengan beberapa jenis serangan kritis, seperti worm dan shellcode, hanya menyumbang kurang dari 0,2% dari total. Ketidakseimbangan ini menyebabkan model deep learning cenderung bias terhadap kelas mayoritas. Li et al. [4] Lebih lanjut disebutkan bahwa pendekatan tradisional untuk mengatasi ketidakseimbangan kelas, seperti synthetic minority oversampling technique (SMOTE), sering menghasilkan sampel sintetis melalui interpolasi linier, yang mungkin gagal menangkap karakteristik non-linier dan kompleks dari data lalu lintas jaringan. Metode ini dapat memperkenalkan pola yang tidak mencerminkan perilaku serangan yang sebenarnya, berpotensi menyebabkan overfitting dan generalisasi data yang buruk.

Di sisi lain, autoencoder variational (VAE)

menawarkan alternatif yang menjanjikan untuk mengatasi masalah data yang tidak seimbang. Tidak seperti teknik oversampling tradisional lainnya, Xie et al. [5] menyatakan bahwa VAE adalah model generatif berbasis deep learning yang menggabungkan prinsip-prinsip autoencoder tradisional dengan inferensi variasional Bayesian untuk mempelajari distribusi laten dari data. Tahir et al. [6] Ia melanjutkan bahwa meskipun studi mereka telah mengeksplorasi VAE untuk deteksi anomali dalam keamanan jaringan, penelitian yang secara sistematis mengevaluasi efektivitas VAE (tanpa arsitektur tambahan) untuk mengatasi ketidakseimbangan kelas dalam IDS masih sangat terbatas, terutama untuk kelas minoritas ekstrem dalam dataset UNSW-NB15.

Topik dan Batasannya

Berdasarkan latar belakang diatas maka permasalahan utama yang diangkat dalam penelitian ini berfokus pada lima aspek. Pertama, bagaimana pengaruh ketidakseimbangan data terhadap performa model dalam mendeteksi serangan worm pada dataset UNSW-NB15. Kedua, bagaimana efektivitas metode Mean Decrease Impurity (MDI) dalam melakukan seleksi fitur dibandingkan dengan Random Forest (RF). Ketiga, Bagaimana kinerja model CatBoost dalam mendeteksi serangan worm dibandingkan dengan AdaBoost. Keempat, apakah penggunaan Variational Autoencoder (VAE) mampu meningkatkan kualitas data minoritas dan performa model klasifikasi. Dan kelima, bagaimana kombinasi metode MDI dan VAE dalam mengatasi masalah ketidakseimbangan data pada IDS. Batasan masalah dalam penelitian ini meliputi dataset yang digunakan adalah UNSW-NB15, penelitian difokuskan pada deteksi serangan worm sebagai kelas minoritas, metode seleksi fitur yang digunakan adalah MDI berbasis Extra Trees Classifier dan dibandingkan dengan Random Forest, model klasifikasi yang digunakan adalah CatBoost dan dibandingkan dengan AdaBoost, teknik penyeimbangan data menggunakan Variational Autoencoder (VAE), Evaluasi kinerja model menggunakan metrik akurasi, presisi, recall, F1-score, AUC, ROC, sensitivitas, dan G-mean, dan penelitian tidak membahas implementasi sistem secara real-time pada jaringan nyata.

Tujuan

Tujuan utama dari penelitian ini adalah menganalisis dampak ketidakseimbangan data terhadap performa deteksi serangan worm, mengevaluasi efektivitas metode MDI dalam seleksi fitur dibandingkan metode Random Forest, Mengkaji performa model CatBoost dalam klasifikasi serangan worm, menguji kemampuan VAE dalam menghasilkan data sintesis untuk mengatasi ketidakseimbangan data, dan mengintegrasikan metode MDI dan VAE untuk meningkatkan kinerja sistem IDS.

Organisasi Tulisan

Sistematika penulisan jurnal ini disusun dalam berbagai tahapan, bagian kedua membahas studi terkait yang menjadi landasan teori dan penelitian terdahulu, serta pada bagian ketiga menjelaskan metodologi penelitian dan alur penelitian yang akan dibangun, keempat memaparkan hasil eksperimen yang mencakup evaluasi dan analisis, dan kelima menyajikan kesimpulan yang

diperoleh dari penelitian serta saran untuk pengembangan dimasa yang akan datang.

2. Studi Terkait

Penelitian terkait Intrusion Detection System (IDS), sangat berkembang pesat dalam beberapa tahun terakhir, seiring meningkatnya keberagaman serangan siber dan kebutuhan untuk mendeteksi anomali pada jaringan yang modern. Dataset *UNSW-NB15* banyak digunakan sebagai patokan atau acuan karena representasinya terhadap lalu lintas jaringan yang nyata. Berikut merupakan beberapa studi yang telah mengeksplorasi IDS, Imbalance dan VAE.

Tabel 1. Tabel keterkaitan antara tujuan, pengujian dan kesimpulan

Reverensi	IDS	Imbalance	VAE
Abbas et al. [7]	✓	✗	✗
Hassounh et al. [8]	✓	✗	✗
Zeeshan et al. [9]	✗	✓	✗
Zoghi et al. [10]	✗	✓	✗
Lin et al. [11]	✗	✗	✓
Tian et al. [12]	✗	✗	✓
Metode Yang diusulkan	✓	✓	✓

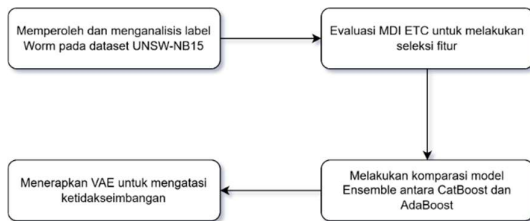
Abbas et al. [7] menyatakan bahwa IDS adalah mekanisme keamanan yang dirancang untuk mengidentifikasi aktivitas yang tidak sah atau mencurigakan dalam infrastruktur jaringan komputer dengan memantau lalu lintas jaringan secara real-time. Lebih lanjut, Hassounh et al. [8] Mereka menerapkan jaringan long short-term memory (LSTM) pada dataset UNSW-NB15, dan mencapai akurasi tinggi untuk klasifikasi biner. Namun, pendekatan mereka tidak secara spesifik mengatasi tantangan ketidakseimbangan kelas. Zeeshan et al. [9] Mengatasi serangan DoS dan DDoS dengan menggabungkan dua dataset untuk membuat batch yang seimbang, mencapai akurasi lebih dari 99% untuk klasifikasi biner tetapi menunjukkan kebingungan antara kedua dataset tersebut. Zoghi et al. [10] Penelitian ini mengembangkan pendekatan ensemble menggunakan balanced bagging, extreme gradient boosting (XGBoost), dan random forests, yang mencapai akurasi 97,80% untuk klasifikasi biner. VAE adalah model generatif berbasis pembelajaran mendalam yang telah mendapat perhatian signifikan karena kemampuannya untuk mempelajari distribusi data laten. Lin et al. [11] Mengaplikasikan VAE untuk menangani masalah imbalanced pada dataset intrusion detection, dengan mengkombinasikan VAE dan Multi-Layer Perceptron (MLP) serta Range-based Sequential Search (RSS), Hasilnya mencapai F1-score 97% dan recall 98% pada log sistem HDFS. Penelitian ini menunjukkan peningkatan skor F1 dibandingkan dengan dataset yang tidak seimbang, menjadikan VAE sebagai pendekatan yang menjanjikan untuk aplikasi keamanan. Tian et al. [12] VAE-WACGAN dikembangkan, yang mengintegrasikan VAE-generative adversarial network (VAEGAN) dengan Auxiliary Classifier Generative Adversarial Network (ACGAN) dan Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) untuk mengatasi ketidakseimbangan kelas melalui pemodelan generatif, mencapai akurasi 0,8579 dan skor F1 0,8552 pada UNSW-NB15. Meskipun kinerjanya baik,

arsitektur hibrida ini menghasilkan kompleksitas dan biaya komputasi yang tinggi. Terdapat peluang untuk menerapkan VAE pada pembuatan data sintesis pada dataset UNSW-NB15

3. Sistem yang Dibangun

3.1. Alur Penelitian

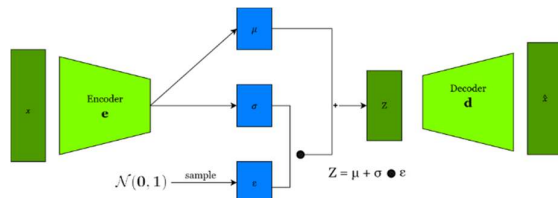
Studi ini mengembangkan alur kerja penelitian untuk mencapai tujuannya. Pada langkah pertama, kami memperoleh label worm dari dataset UNSW-NB15. Pada langkah kedua dan ketiga, kami menerapkan MDI ETC untuk seleksi fitur dan CatBoost untuk deteksi label worm. Langkah terakhir adalah mengevaluasi AUC VAE selama pembuatan data sintesis untuk mengatasi ketidakseimbangan label worm dalam dataset UNSW-NB15.



Gambar 1 Diagram Alur Penelitian

3.2. VAE untuk Generasi Sample Sintetis

Studi ini menggunakan VAE untuk mempelajari distribusi laten lalu lintas serangan worm dan untuk menghasilkan sampel sintesis.



Gambar 2 Arsitektur VAE

Pada gambar 2 menunjukkan arsitektur VAE. Arsitektur VAE terdiri dari beberapa komponen utama. Pertama adalah jaringan encoder, yang memetakan vektor fitur input ke ruang laten probabilistik [13]. Rumusnya adalah sebagai berikut:

$$f(\phi, z, x) = q_{\phi}(z | x) = \mathcal{N}\left(z; \mu_{\phi}(x), \sigma_{\phi}^2(x)\right)$$

Dimana $q_{\phi}(z | x)$ adalah distribusi data masukan x yang direkonstruksi dari variabel laten z , $\mathcal{N}$ adalah distribusi normal atau Gaussian, $\mu_{\phi}(x)$ = fungsi rata-rata yang dihasilkan oleh encoder dengan parameter ϕ dan $\sigma_{\phi}^2(x)$ = fungsi varians yang dihasilkan oleh encoder dengan parameter ϕ . Bagian kedua adalah jaringan dekoder, yang merekonstruksi vektor fitur dari representasi laten z . Arsitektur dekoder bersifat simetris dan ringan, sehingga menjamin stabilitas pada dataset besar. Rumusnya adalah $f(\phi, z, x) = p_{\theta}(x | z)$ dimana $p_{\theta}(x | z)$ adalah distribusi kemungkinan, dan θ adalah parameter jaringan dekoder.

Selanjutnya, bagian berikutnya adalah fungsi kerugian, di mana VAE dilatih dengan meminimalkan kerugian L_{VAE} , yang merupakan akumulasi dari reconstruction loss L_{recon} dan Kullback-Leibler (KL) L_{KL} . [14]

Rumusnya adalah sebagai berikut:

$$L_{VAE} = L_{recon} + \beta L_{KL}$$

Dimana β hyperparameter bobot untuk menyeimbangkan kedua komponen loss.

Dalam prosedur pelatihan, VAE dilatih hanya pada data serangan Worms dari set pelatihan untuk menghindari kebocoran distribusi kelas normal. Setelah pelatihan selesai, decoder menghasilkan sampel sintesis dengan mengambil sampel dari vektor laten. Sampel worm sintesis ini kemudian digabungkan dengan data pelatihan asli untuk membentuk set pelatihan yang lebih seimbang, sementara data validasi dan pengujian tetap tidak berubah [15].

Setelah pelatihan VAE berhasil, sampel serangan worm sintesis dihasilkan untuk mengurangi ketidakseimbangan kelas dalam data pelatihan. Proses pembuatan sampel sintesis melibatkan beberapa tahapan. Tahap pertama adalah pengambilan sampel laten, di mana vektor laten sintesis diambil dari prior Gaussian standar yang dipelajari oleh VAE. Rumusnya adalah sebagai berikut [13]:

$$z_{syn} \sim \mathcal{N}(0,1)$$

Dimana z_{syn} adalah vektor laten sintesis dan $\mathcal{N}(0,1)$ adalah distribusi normal standar.

Pada tahap decoding, vektor laten sintesis dilewatkan melalui decoder yang telah dilatih untuk menghasilkan vektor fitur sintesis. Rumusnya adalah sebagai berikut:

$$x_{syn} = Decoder(z_{syn})$$

Dimana x_{syn} adalah vektor fitur sintesis yang telah didekode, dan $(.)$ adalah fungsi jaringan decoder.

Pembuatan sampel sintesis menghasilkan sampel worm sintesis hingga himpunan pelatihan yang diperluas mencapai 80% dari kelas normal, menyeimbangkan kinerja deteksi dengan efisiensi pelatihan. Himpunan pelatihan akhir menggabungkan sampel nyata dan sintesis, dan pengacakan acak memastikan distribusi yang seragam selama pelatihan batch

3.3. Model Ensemble CatBoost dengan Pemilihan Fitur MDI-ETC

One-hot encoding adalah metode yang mengubah fitur kategorikal menjadi fitur numerik dengan membuat fitur baru berdasarkan jumlah nilai unik dalam suatu fitur dan menetapkan nilai biner untuk setiap fitur baru. Namun, onehot encoding dapat menyebabkan kutukan dimensi jika fitur memiliki kardinalitas tinggi. Studi ini menerapkan seleksi fitur MDI untuk mencegah hal ini. Seleksi fitur MDI terkini untuk UNSW NB15 menggunakan Gini impurity dari random forest. ETC, di sisi lain, memperluas random forest dengan menggunakan pemisahan acak, sehingga mengurangi varians prediksi [16].

Studi ini menggunakan CatBoost untuk mendeteksi serangan worm pada dataset UNSW-NB15. CatBoost adalah metode pembelajaran ensemble yang memperluas gradient boosting dengan mengatasi kebocoran target dan pergeseran prediksi. Hal ini dilakukan dengan menerapkan boosting terurut, di mana model pembelajar lemah yang telah berpartisipasi dalam pelatihan tidak lagi disertakan dalam tahap pelatihan selanjutnya [17]. Metode ini dapat mengatasi pembelajaran yang terlalu optimis dan overfitting. Penelitian ini membandingkan CatBoost dengan pembelajaran ensemble terkini untuk UNSW-NB15: AdaBoost. Algoritma 1 Menunjukkan algoritma untuk

boosting terurut di CatBoost.

Algorithm 1: The algorithm for ordered boosting in CatBoost

Input Training Data

Output Trained CatBoost Model

```

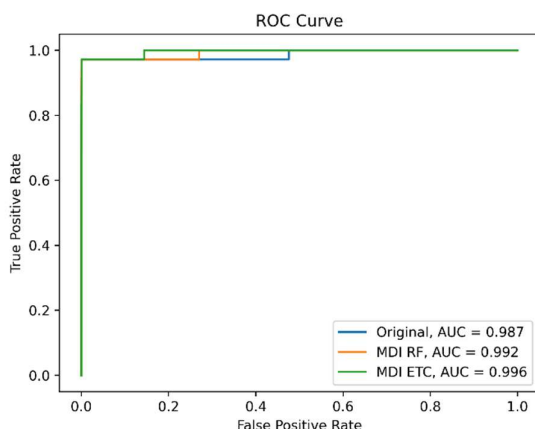
1 Randomly permute training data
2 for each position  $i$  do
3   Train a temporary model using 1, ...,  $i-1$ 
4   Predict sample  $i$ 
5   Compute the residual
6 Use these clean residuals to grow trees
  
```

4. Evaluasi

Penelitian ini memulai fase pengujian dengan memperoleh dataset open-source UNSW-NB15. Dari dataset ini, kami menghapus item data dengan label selain worm dan normal. Dataset UNSW-NB15 yang tersisa menunjukkan ketidakseimbangan kelas yang ekstrem dalam skenario klasifikasi biner. Pada tahap selanjutnya, penelitian ini menerapkan onehot encoding pada dataset untuk fitur kategorikal protocol, service, dan state. Setelah tahap itu, jumlah fitur berubah dari 44 menjadi 66. Kami menggunakan seleksi fitur MDI ETC untuk mengurangi jumlah fitur. Selanjutnya, kami membandingkan MDI ETC dengan metode terkini, MDI RF. Kami menggunakan akurasi, presisi, recall, dan F1-score sebagai metrik perbandingan. Nilai presisi yang lebih rendah daripada nilai recall pada kedua

4.1 Hasil Feature Selection

metode pemilihan fitur menunjukkan ketidakseimbangan dalam dataset. Berdasarkan presisi dan F1-score, penelitian ini menemukan bahwa MDI ETC mengungguli MDI RF. Presisi dan F1-score MDI ETC masing-masing adalah 0,80 dan 0,87. Sementara itu, presisi dan F1-score MDI RF masing-masing adalah 0,61 dan 0,68. Kami menggunakan MDI ETC untuk langkah optimasi selanjutnya.



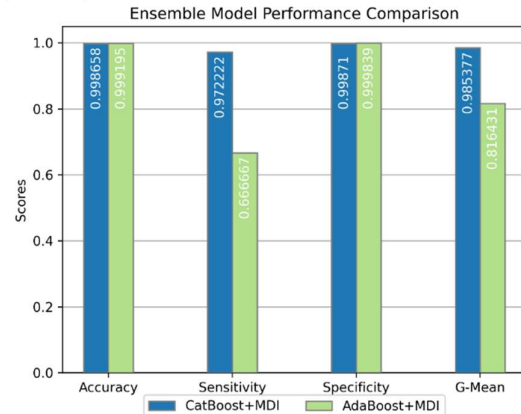
Gambar 3 Performa deteksi serangan worm berdasarkan metode pemilihan fiturnya.

Gambar tersebut menunjukkan perbandingan kurva ROC dari kedua metode, di mana MDI ETC terbukti memiliki AUC terbesar dibandingkan dengan MDI RF dan model tanpa pemilihan fitur.

4.2 Hasil Model Ensemble

Dalam pengujian berikut, kami membandingkan dua

metode ensemble: ensemble yang kami usulkan, CatBoost, dan ensemble terkini, AdaBoost. Studi ini menggunakan tiga metrik tambahan: sensitivity, specificity, and G-mean.



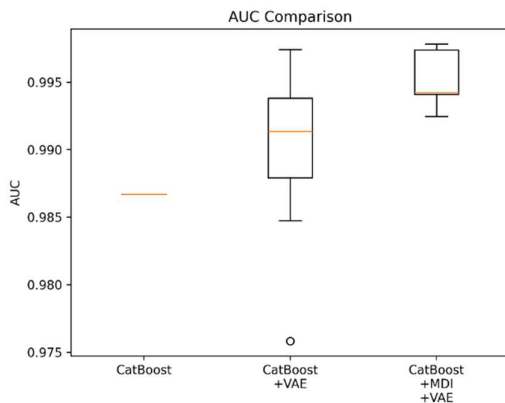
Gambar 4 Perbandingan kinerja model ensemble menggunakan akurasi, sensitivitas, spesifisitas, dan G-means

Gambar tersebut menunjukkan perbandingan kinerja kedua metode ensemble. Hasil penelitian ini menunjukkan bahwa meskipun akurasi dan spesifisitas AdaBoost+MDI melebihi CatBoost, sensitivitas dan G-mean-nya lebih rendah daripada CatBoost. Selain itu, G-mean adalah metrik yang menggabungkan sensitivitas dan spesifisitas. Oleh karena itu, CatBoost+MDI adalah metode yang lebih unggul untuk deteksi worm dengan skor sensitivitas dan G-mean masing-masing sebesar 0,97 dan 0,99.

Dalam set pelatihan, terdapat 56.000 label normal (99,77%), 130 label worm (0,23%), dengan total 56.130. Rasio Ketidakseimbangan (IR) adalah 431 kali. Dalam dataset pengujian, terdapat 37.000 label normal (99,88%) dan 44 label worm (0,12%), dengan total 37.044. IR adalah 841 kali. Ketidakseimbangan ini menimbulkan masalah yang menantang bagi model klasifikasi, terutama mengingat jumlah sampel kelas minoritas yang terbatas dalam set pengujian.

4.3 Hasil Pengujian VAE

Sampai saat ini, kami telah menguji VAE pada dua model: CatBoost saja dan CatBoost+MDI; oleh karena itu, fase pengujian ini membandingkan ketiga model tersebut. Studi ini membandingkan AUC dari ketiga model menggunakan boxplot berdasarkan 10 pengujian



Gambar 5 Perbandingan boxplot dari tiga metodologi gabungan AUC untuk ketidakseimbangan dalam dataset UNSW-NB15, deteksi label Worm

Gambar tersebut menunjukkan perbandingan boxplot dari ketiga model. Hasil pengujian menunjukkan bahwa CatBoost saja, CatBoost+VAE, dan CatBoost+MDI+VAE memiliki VAE rata-rata masing-masing sebesar 0,987, 0,990, dan 0,995. Meskipun memiliki AUC rata-rata terendah, CatBoost menunjukkan kekokohan terendah, dengan deviasi standar masing-masing sebesar 0,0, 0,006, dan 0,002. Studi ini juga menguji signifikansi perbedaan di antara ketiga AUC; hasilnya menunjukkan tidak ada perbedaan signifikan dalam AUC rata-rata antara CatBoost saja dan CatBoost+VAE. Namun, nilai p antara CatBoost saja dan CatBoost+VAE dengan CatBoost+MDI+VAE adalah nilai $p = 0,01$ dan nilai $p = 0,02$, sehingga MDI+VAE memiliki pengaruh signifikan terhadap AUC CatBoost, dan MDI memiliki pengaruh signifikan terhadap AUC CatBoost+VAE.

4.4. Evaluasi

Bouke et al. [18] menggunakan MDI RF untuk feature selection pada dataset UNSW-NB-15. Hornung et al. [19] menyatakan bahwa ETC menggunakan pembagian acak selama pelatihan dan metode ini memberikan ketahanan selama pengujian. Tes kami menunjukkan bahwa pemilihan fitur berdasarkan MDI ETC menghasilkan kinerja yang lebih baik daripada MDI RF pada CatBoost. Ini menunjukkan bahwa pembagian acak oleh MDI ETC meningkatkan pemilihan fitur pada dataset UNSWNB15. Kontribusi penelitian kami adalah metode pemilihan fitur optimal berdasarkan MDI ETC yang tangguh karena random splitting.

Fauzi et al. [20] menyatakan bahwa model ensemble merupakan metode yang kuat untuk IDS menggunakan UNSWNB15. Studi ini mengusulkan penggunaan pengklasifikasi AdaBoost sebagai metode ensemble-nya. Meanwhile, Filom et al. [21] menyatakan bahwa CatBoost dapat mengungguli AdaBoost karena dapat menangkap fitur interaksi melalui pohon simetris. Dalam penelitian kami, CatBoost mengungguli AdaBoost, mencapai G-mean 0,985 dibandingkan 0,816. Kontribusi penelitian ini adalah model ensemble optimal untuk deteksi serangan worm pada dataset UNSW-NB15 menggunakan CatBoost yang menangkap fitur interaksi melalui model simetris.

Al Amien et al. [22] menyatakan bahwa beberapa label di UNSW-NB15 sangat tidak seimbang. Studi kami menunjukkan bahwa IR label worm adalah $431 \times$ untuk data pelatihan dan $841 \times$ untuk data uji. Di sisi lain, Wang et al. [23] Telah dilaporkan bahwa VAE dapat mengatasi dataset yang tidak seimbang, seperti yang ditunjukkan dalam studi kasus tentang klasifikasi cacat wafer. Namun, belum ada penelitian yang menggunakan VAE untuk mengatasi distribusi label worm yang tidak seimbang dalam dataset UNSW-NB15. Pengujian kami menunjukkan bahwa VAE tidak secara signifikan meningkatkan AUC CatBoost, sedangkan MDI+VAE meningkatkan AUC rata-rata dari 0,987 menjadi 0,995. Kontribusi penelitian kami adalah model pembangkitan data sintesis optimal untuk mengatasi label worm yang tidak seimbang di UNSW-NB15 menggunakan MDI+VAE.

5. Kesimpulan

Studi ini menyajikan evaluasi empiris komprehensif dari VAE sebagai pendekatan metodologis untuk mengatasi ketidakseimbangan kelas ekstrem dalam sistem deteksi intrusi jaringan, dengan aplikasi khusus pada label worm dalam dataset UNSW-NB15. Kami menggunakan seleksi fitur, yaitu MDI ETC, untuk mengurangi jumlah fitur dalam UNSWNB15. Studi ini juga menggunakan CatBoost sebagai metode ensemble untuk mendeteksi label worm.

Hasil pengujian menunjukkan bahwa CatBoost, menggunakan MDI ETC, mengungguli MDI RF, mencapai presisi dan skor G-mean masing-masing 0,80 dan 0,86, dibandingkan dengan 0,61 dan 0,68. CatBoost mengungguli AdaBoost dalam sensitivitas dan G mean, mencapai 0,97 dan 0,99, dibandingkan dengan 0,67 dan 0,82. Terakhir, penerapan VAE tidak memiliki efek signifikan pada AUC CatBoost, sedangkan MDI+VAE meningkatkannya dari 0,987 menjadi 0,995. Pekerjaan selanjutnya dapat difokuskan pada implementasi deteksi worm di jaringan IoT menggunakan teknik kompresi model, seperti pemangkasan, untuk ETC.

Daftar Pustaka

- [1] M. S. Al-Daweri, K. A. Zainol Ariffin, S. Abdullah, and M. F. E. Md. Senan, "An Analysis of the KDD99 and UNSW-NB15 Datasets for the Intrusion Detection System," *Symmetry*, vol. 12, no. 10, p. 1666, Oct. 2020, doi: 10.3390/sym12101666.
- [2] S. W. Fernandes, A. Gautama Putrada, R. L. Wicaksono, and F. Dea Reskyadita, "Tomek Links and Random Undersampling for Infiltration Attack Detection in Imbalanced CSE-CIC-IDS2018 Dataset," in *2025 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, Padang, Indonesia: IEEE, Dec. 2025, pp. 272–277. doi: 10.1109/COMNETSAT68601.2025.11324919.
- [3] H. Xie, Y. Shao, Z. Li, Z. Alomari, and A. Makanju, "Optimization of Class Imbalance Techniques in Machine Learning Models for Network Intrusion Detection," in *2025 9th International Conference on Cryptography, Security and Privacy (CSP)*, Okinawa, Japan: IEEE, Apr. 2025, pp. 102–106. doi: 10.1109/CSP66295.2025.00025.

- [4] M. Li, H. Zhou, Q. Liu, Y. Shao, and G. Wang, "INGB: Informed Nonlinear Granular Ball Oversampling Framework for Noisy Imbalanced Classification," 2023, *arXiv*. doi: 10.48550/ARXIV.2307.01224.
- [5] B. Xie, X. Xu, and G. Wen, "Network Intrusion Detection Optimization based on Graph Neural Networks and Variational Autoencoders," in *2024 6th International Conference on Frontier Technologies of Information and Computer (ICFTIC)*, Qingdao, China: IEEE, Dec. 2024, pp. 127–134. doi: 10.1109/ICFTIC64248.2024.10912964.
- [6] S. Tahir, "Variational Autoencoders (VAEs) for Anomaly Detection:," in *Advances in Digital Crime, Forensics, and Cyber Terrorism*, N. Z. Jhanjhi, Ed., IGI Global, 2024, pp. 309–326. doi: 10.4018/979-8-3693-8944-7.ch009.
- [7] Safana Hyder Abbas, Wedad Abdul Khuder Naser, and Amal Abbas Kadhim, "Subject review: Intrusion Detection System (IDS) and Intrusion Prevention System (IPS)," *Glob. J. Eng. Technol. Adv.*, vol. 14, no. 2, pp. 155–158, Feb. 2023, doi: 10.30574/gjeta.2023.14.2.0031.
- [8] N. Hassouneh and S. Al-sharaeh, "Intrusion Detection in IoT Networks Using LSTM Deep Learning Models with the UNSW-NB15 Dataset," in *2025 International Conference on New Trends in Computing Sciences (ICTCS)*, Amman, Jordan: IEEE, Apr. 2025, pp. 263–269. doi: 10.1109/ICTCS65341.2025.10989358.
- [9] M. Zeeshan *et al.*, "Protocol-Based Deep Intrusion Detection for DoS and DDoS Attacks Using UNSW-NB15 and Bot-IoT Data-Sets," *IEEE Access*, vol. 10, pp. 2269–2283, 2022, doi: 10.1109/ACCESS.2021.3137201.
- [10] Z. Zoghi and G. Serpen, "Building an intrusion detection system on UNSW - NB15: Reducing the margin of error to deal with data overlap and imbalance," *Concurr. Comput. Pract. Exp.*, vol. 36, no. 25, p. e8242, Nov. 2024, doi: 10.1002/cpe.8242.
- [11] Y.-D. Lin, Z.-Q. Liu, R.-H. Hwang, V.-L. Nguyen, P.-C. Lin, and Y.-C. Lai, "Machine Learning With Variational AutoEncoder for Imbalanced Datasets in Intrusion Detection," *IEEE Access*, vol. 10, pp. 15247–15260, 2022, doi: 10.1109/ACCESS.2022.3149295.
- [12] W. Tian, Y. Shen, N. Guo, J. Yuan, and Y. Yang, "VAE-WACGAN: An Improved Data Augmentation Method Based on VAEGAN for Intrusion Detection," *Sensors*, vol. 24, no. 18, p. 6035, Sep. 2024, doi: 10.3390/s24186035.
- [13] D. G. Edelberg and R. R. Lederman, "Using VAEs to Learn Latent Variables: Observations on Applications in cryo-EM," 2023, *arXiv*. doi: 10.48550/ARXIV.2303.07487.
- [14] "1244_A_prescriptive_theory_for."
- [15] S. Dodda, "Exploring Variational Autoencoders and Generative Latent Time-Series Models for Synthetic Data Generation and Forecasting," in *2024 Control Instrumentation System Conference (CISCON)*, Manipal, India: IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/CISCON62171.2024.10696588.
- [16] G. Alfian *et al.*, "Predicting Breast Cancer from Risk Factors Using SVM and Extra-Trees-Based Feature Selection Method," *Computers*, vol. 11, no. 9, p. 136, Sep. 2022, doi: 10.3390/computers11090136.
- [17] S. Geeitha, K. Ravishankar, J. Cho, and S. V. Easwaramoorthy, "Integrating cat boost algorithm with triangulating feature importance to predict survival outcome in recurrent cervical cancer," *Sci. Rep.*, vol. 14, no. 1, p. 19828, Aug. 2024, doi: 10.1038/s41598-024-67562-0.
- [18] M. A. Bouke, A. Abdullah, J. Frnda, K. Cengiz, and B. Salah, "BukaGini: A Stability-Aware Gini Index Feature Selection Algorithm for Robust Model Performance," *IEEE Access*, vol. 11, pp. 59386–59396, 2023, doi: 10.1109/ACCESS.2023.3284975.
- [19] R. Hornung, "Diversity Forests: Using Split Sampling to Enable Innovative Complex Split Procedures in Random Forests," *SN Comput. Sci.*, vol. 3, no. 1, p. 1, Jan. 2022, doi: 10.1007/s42979-021-00920-1.
- [20] N. Fauzi, F. A. Yulianto, and H. H. Nuha, "The Effectiveness of Anomaly-Based Intrusion Detection Systems in Handling Zero-Day Attacks Using AdaBoost, J48, and Random Forest Methods," in *2023 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bali, Indonesia: IEEE, Oct. 2023, pp. 57–62. doi: 10.1109/APWiMob59963.2023.10365642.
- [21] K. Filom, A. Miroshnikov, K. Kotsiopoulos, and A. R. Kannan, "On marginal feature attributions of tree-based models," *Found. Data Sci.*, vol. 6, no. 4, pp. 395–467, 2024, doi: 10.3934/fods.2024021.
- [22] J. A. Amien, Yoze Rizki, and Mukhlis Ali Rahman Nasution, "Implementasi Adasyn Untuk Imbalance Data Pada Dataset UNSW-NB15 Adasyn Implementation For Data Imbalance on UNSW-NB15 Dataset," *J. CoSciTech Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 242–248, Dec. 2022, doi: 10.37859/coscitech.v3i3.4339.
- [23] Y. Wang, Y. Wei, and H. Wang, "A class imbalanced wafer defect classification framework based on variational autoencoder generative adversarial network," *Meas. Sci. Technol.*, vol. 34, no. 2, p. 024008, Feb. 2023, doi: 10.1088/1361-6501/ac9ed3.