**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
August 2026, IX (4): 6124 – 6132 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

# ANALISIS SELEKSI FITUR PADA *NETWORK INTRUSION DETECTION* DENGAN NATURAL *LANGUAGE PROCESSING*

**Riki Andri Yusda[1], Sahren[2], Nurwati[3], Shaqilla Swita Sarah[4], Dava Erlangga[5]**
**Universitas Royal, Kisaran**
e-mail: [1]rikiandriyusda@gmail.com, [2]sahren.one@gmail.com, [3]nurwati763@gmail.com

***Abstract:*** *This study proposes a Natural Language Processing (NLP)-based feature selection approach for Network Intrusion Detection Systems (NIDS) using the CICIDS2017 dataset. To address high computational complexity caused by high dimensionality, network flows are normalized using Min-Max Scaling and transformed into textual representations via a consistent sentence template. The text data is then weighted using Term Frequency–Inverse Document Frequency (TF-IDF), followed by Chi-Square (Chi²) feature selection to isolate the most discriminative features. This process yields a dataset comprising 691,406 text documents with 302 textual features. Performance is evaluated using Random Forest, Linear Support Vector Machine (Linear SVM), and Naïve Bayes classifiers. Experimental results show that Random Forest achieves the highest accuracy at 97.3%, followed by Linear SVM at 96.4%, and Naïve Bayes at 83.2%. The confusion matrix further confirms that the vast majority of network traffic instances are correctly classified. These findings demonstrate that transforming tabular network data into textual formats using NLP, combined with TF-IDF and Chi-Square feature selection, successfully produces representative features, enhances intrusion detection performance, and effectively reduces data complexity in NIDS.*

***Keywords:*** *NIDS; NLP; TF-IDF; Chi-Square; Feature Selection.*

**Abstrak:** Meningkatnya kompleksitas serangan siber menuntut Network Intrusion Detection System (NIDS) memiliki kemampuan deteksi yang akurat dan efisien. Salah satu tantangan utama dalam pengembangan NIDS adalah tingginya dimensi fitur pada data lalu lintas jaringan yang dapat meningkatkan kompleksitas komputasi dan memengaruhi kinerja model klasifikasi. Penelitian ini mengusulkan pendekatan berbasis Natural Language Processing (NLP) untuk melakukan seleksi fitur pada dataset CICIDS2017. Data terlebih dahulu dinormalisasi menggunakan Min-Max Scaling, kemudian setiap network flow ditransformasikan menjadi representasi bahasa alami melalui template kalimat yang konsisten. Representasi teks selanjutnya dibobot menggunakan Term Frequency–Inverse Document Frequency (TF-IDF) dan dilakukan seleksi fitur menggunakan metode Chi-Square (Chi²) untuk memperoleh fitur yang paling relevan. Dataset hasil transformasi terdiri atas 691.406 dokumen dengan representasi vektor berdimensi 302 fitur tekstual. Evaluasi dilakukan menggunakan algoritma Random Forest, Linear Support Vector Machine, dan Naive Bayes. Hasil pengujian menunjukkan bahwa Random Forest memberikan performa terbaik dengan akurasi sebesar 97,3%, Linear SVM sebesar 96,4%, Naive Bayes memperoleh akurasi 83,2%. Hasil confusion matrix juga menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar. Penelitian menunjukkan bahwa transformasi data tabular menjadi representasi teks menggunakan NLP yang dipadukan dengan TF-IDF dan Chi-Square mampu menghasilkan fitur yang lebih representatif, meningkatkan kinerja deteksi intrusi, serta mengurangi kompleksitas data pada NIDS.

**Kata kunci:** NIDS; NLP; TF-IDF; Chi-Square; Seleksi Fitur.

## PENDAHULUAN

Perkembangan teknologi informasi

telah meningkatkan konektivitas jaringan pada berbagai sektor, seperti pemerintahan, pendidikan, kesehatan, industri, dan layanan keuangan. Di sisi lain, peningkatan konektivitas tersebut juga memperbesar risiko terjadinya serangan siber, seperti *Distributed Denial of Service* (DDoS), *SQL Injection*, *brute force*, *malware*, *botnet*, hingga *zero-day attack* (Cybersecurity Ventures, 2023) (Ponemon Institute, 2024) (Sahren & Prijuna Lubis, 2024). Kondisi ini menjadikan Network Intrusion Detection System (NIDS) sebagai salah satu komponen penting dalam menjaga keamanan jaringan melalui kemampuan mendeteksi aktivitas yang menyimpang secara dini (Talukder et al., 2023). Berbagai pendekatan berbasis *Machine Learning* (ML) dan *Deep Learning* (DL) telah berhasil meningkatkan kemampuan deteksi dibandingkan dengan metode berbasis aturan (*rule-based*), namun performa model masih sangat dipengaruhi oleh kualitas fitur yang digunakan. Banyaknya fitur yang tidak relevan maupun redundan menyebabkan peningkatan kompleksitas komputasi, penggunaan memori yang lebih besar, serta penurunan akurasi klasifikasi akibat munculnya *noise* pada data. Oleh karena itu, proses seleksi fitur menjadi tahapan penting untuk menghasilkan model NIDS yang lebih efisien dan akurat (Sworna et al., 2023) (Yin et al., 2023).

Berbagai penelitian terbaru telah mengembangkan metode seleksi fitur untuk meningkatkan performa NIDS. Ngo dkk. (2023) menunjukkan bahwa pendekatan *feature selection* mampu memberikan waktu komputasi yang lebih rendah dibandingkan dengan *feature extraction* tanpa mengorbankan akurasi deteksi pada dataset modern (Ngo et al., 2023). Selanjutnya, Westphal dkk. (2025) mengusulkan metode *Feature Selection for Network Intrusion Detection* (FSNID) berbasis teori informasi yang mampu mengurangi jumlah fitur secara signifikan dengan tetap mempertahankan performa deteksi (Westphal et al., 2024). Selain itu, penelitian Arreche dkk. (2024) memanfaatkan pendekatan Explainable Artificial Intelligence (XAI) untuk mengidentifikasi atribut yang paling berpengaruh terhadap keputusan model IDS sehingga meningkatkan interpretabilitas proses seleksi fitur (Arreche et al., 2024). Meskipun demikian, penelitian-penelitian tersebut masih berorientasi pada karakteristik statistik maupun numerik data jaringan sehingga belum mempertimbangkan makna semantik dari atribut-atribut yang digunakan dalam proses seleksi fitur.

Di sisi lain, perkembangan *Natural Language Processing* (NLP), khususnya melalui model *Bidirectional Encoder Representations from Transformers* (BERT), telah membuka peluang baru dalam analisis keamanan siber. (Yang & Peng, 2025) membuktikan bahwa BERT mampu meningkatkan performa NIDS dengan mengubah data lalu lintas jaringan menjadi representasi berbasis teks sehingga hubungan kontekstual antar data dapat dipelajari secara lebih baik. Penelitian serupa juga dilakukan oleh (Maasaoui et al., 2024)yang memanfaatkan *Large Language Models* berbasis BERT untuk mendeteksi anomali pada lingkungan IoT dan memperoleh peningkatan akurasi dibandingkan pendekatan konvensional. Meskipun demikian, kedua penelitian tersebut menggunakan BERT sebagai *feature extractor* maupun model klasifikasi, bukan sebagai metode untuk melakukan seleksi fitur pada atribut jaringan. Dengan demikian, potensi BERT dalam memahami hubungan semantik antar fitur jaringan masih belum dimanfaatkan secara optimal (Vubangsi et al., 2024) (Mamikonian & Mamamniashvili, 2025) (Umer et al., 2025) (Ma et al., 2025).

Berdasarkan penelitian-penelitian tersebut, dapat diidentifikasi adanya kesenjangan penelitian (*research gap*). Pertama, sebagian besar metode seleksi fitur pada NIDS masih mengandalkan pendekatan statistik, optimasi, atau pentingnya atribut berdasarkan nilai numerik. Kedua, penelitian yang mengintegrasikan BERT pada NIDS lebih

banyak difokuskan pada ekstraksi fitur dan klasifikasi lalu lintas jaringan, bukan pada proses pemilihan fitur yang paling relevan. Ketiga, belum banyak penelitian yang mengeksplorasi pemanfaatan representasi semantik BERT untuk menganalisis hubungan kontekstual antara nama dan deskripsi atribut jaringan sebagai dasar seleksi fitur. Padahal, informasi semantik yang terkandung dalam atribut jaringan berpotensi meningkatkan kemampuan model dalam memilih fitur yang representatif sekaligus mengurangi kompleksitas data masukan.

Berdasarkan permasalahan dan kesenjangan penelitian tersebut, penelitian ini mengusulkan BERT-NIDS, yaitu pendekatan seleksi fitur pada Network Intrusion Detection System yang memanfaatkan representasi semantik dari model BERT untuk mengidentifikasi fitur-fitur jaringan yang paling relevan. Berbeda dengan penelitian sebelumnya yang menggunakan BERT sebagai model klasifikasi atau ekstraksi fitur, penelitian ini menempatkan BERT sebagai dasar dalam proses seleksi fitur sehingga diharapkan mampu menghasilkan subset fitur yang lebih representatif, mengurangi dimensi data, meningkatkan efisiensi komputasi, serta mempertahankan atau bahkan meningkatkan akurasi deteksi intrusi. Hasil penelitian ini diharapkan dapat memberikan kontribusi baru dalam pengembangan metode seleksi fitur berbasis NLP untuk sistem deteksi intrusi jaringan yang lebih adaptif terhadap karakteristik serangan siber modern.

## METODE

Penelitian ini menggunakan metode eksperimen (*experimental research*) untuk mengembangkan dan mengevaluasi pendekatan BERT-NIDS sebagai metode seleksi fitur pada *Network Intrusion Detection System* (NIDS). Tahapan penelitian meliputi pengumpulan *dataset*, *preprocessing data*, representasi semantik fitur menggunakan BERT, proses seleksi fitur, pembangunan model klasifikasi, serta evaluasi kinerja model. Alur penelitian dimulai dari pengolahan data mentah hingga analisis hasil untuk mengetahui efektivitas metode yang diusulkan dibandingkan dengan metode seleksi fitur konvensional.


**Gambar 1 Tahapan Penelitian**

Pada tahap awal, penelitian menggunakan *dataset Network Intrusion Detection* yang banyak digunakan sebagai standar evaluasi, seperti CICIDS2017 (Sharafaldin et al., 2018). Dataset tersebut berisi data lalu lintas jaringan yang telah diberi label normal maupun berbagai jenis serangan siber, sehingga dapat digunakan untuk membangun dan mengevaluasi model deteksi intrusi. Tahap *preprocessing* bertujuan menghasilkan data yang bersih dan siap digunakan dalam proses pembelajaran mesin. Proses ini meliputi penghapusan data duplikat, penanganan *missing value*, transformasi label kategorikal menjadi numerik (*label encoding*), serta normalisasi data agar seluruh fitur memiliki rentang nilai yang seragam. Normalisasi dilakukan menggunakan metode Min-Max Normalization, yang dinyatakan dengan persamaan:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

dengan x merupakan nilai awal suatu fitur, $x_{\min}$ dan $x_{\max}$ adalah nilai minimum dan maksimum pada fitur tersebut, sedangkan x′ merupakan nilai hasil normalisasi. Tahap ini bertujuan

untuk mengurangi perbedaan skala antarfitur sehingga proses pelatihan model menjadi lebih stabil.

Tahap berikutnya merupakan inti dari penelitian, yaitu menghasilkan representasi semantik setiap atribut jaringan menggunakan model BERT. Berbeda dengan metode seleksi fitur konvensional yang hanya memanfaatkan nilai numerik, penelitian ini mengubah nama maupun deskripsi setiap fitur menjadi representasi vektor (*embedding*) yang mampu menggambarkan hubungan kontekstual antarfitur. Secara matematis, proses pembentukan representasi fitur menggunakan BERT dapat dinyatakan sebagai:

$$E = BERT(X)$$

dengan X merupakan kumpulan *token* hasil tokenisasi nama atau deskripsi fitur, sedangkan E merupakan representasi *embedding* yang dihasilkan oleh model BERT. *Embedding* tersebut selanjutnya digunakan sebagai dasar untuk menganalisis hubungan semantik antarfitur.

Setelah diperoleh representasi semantik setiap fitur, dilakukan proses seleksi fitur untuk menentukan atribut yang paling relevan terhadap proses deteksi intrusi. Hubungan antarfitur dianalisis berdasarkan tingkat kemiripan (*semantic similarity*) menggunakan *Cosine Similarity*, karena metode ini mampu mengukur kedekatan dua buah vektor hasil *embedding*. Persamaan *cosine similarity* dinyatakan sebagai berikut:

$$Similarity(f_i, f_j) = \frac{v_i . v_j}{\|v_i\| \|v_j\|}$$

dengan $v_i$ dan $v_j$ merupakan vektor embedding dari dua fitur yang dibandingkan. Nilai yang semakin mendekati 1 menunjukkan bahwa kedua fitur memiliki hubungan semantik yang semakin kuat. Berdasarkan nilai kemiripan tersebut, fitur-fitur kemudian diberi peringkat (feature ranking) dan dipilih sejumlah fitur terbaik (Top-k features) untuk digunakan pada tahap klasifikasi.

Dataset yang telah direduksi menggunakan metode BERT-NIDS selanjutnya digunakan sebagai data masukan untuk model klasifikasi. Penelitian dapat menggunakan algoritma seperti *Random Forest*, *XGBoost*, atau *LightGBM* sebagai pembanding untuk mengetahui pengaruh seleksi fitur terhadap kinerja deteksi intrusi. Penggunaan lebih dari satu algoritma bertujuan untuk mengevaluasi konsistensi metode yang diusulkan pada berbagai model klasifikasi.

Tahap akhir dilakukan dengan membandingkan performa BERT-NIDS dengan metode seleksi fitur konvensional, seperti *Information Gain*, *Chi-Square*, *Mutual Information*, atau *Recursive Feature Elimination* (RFE) (Yusda et al., 2025). Analisis dilakukan berdasarkan jumlah fitur yang dipilih, peningkatan akurasi klasifikasi, efisiensi waktu komputasi, serta kemampuan model dalam mendeteksi berbagai jenis serangan siber. Hasil analisis ini digunakan untuk menilai apakah pendekatan berbasis representasi semantik menggunakan BERT mampu memberikan peningkatan performa dibandingkan dengan metode seleksi fitur yang hanya memanfaatkan karakteristik statistik atau numerik data jaringan.

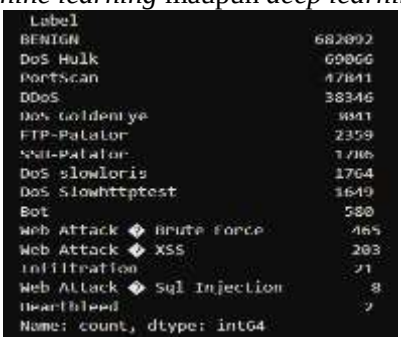
## HASIL DAN PEMBAHASAN

### Dataset CICIDS-2017

Penelitian ini menggunakan *Canadian Institute for Cybersecurity Intrusion Detection System 2017* (CICIDS2017) sebagai sumber data utama dalam proses pengembangan dan evaluasi model NIDS berbasis NLP. Dataset ini dikembangkan oleh Canadian Institute for Cybersecurity (CIC), University of New Brunswick, dengan tujuan menyediakan dataset yang lebih representatif dibandingkan dengan dataset IDS generasi sebelumnya, seperti KDD Cup 1999 dan NSL-KDD. CICIDS2017

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
August 2026, IX (4): 6124 – 6132 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

dirancang untuk merepresentasikan kondisi lalu lintas jaringan yang mendekati lingkungan nyata (*real-world network traffic*), sehingga banyak digunakan sebagai *benchmark* dalam penelitian deteksi intrusi berbasis *machine learning* maupun *deep learning*.

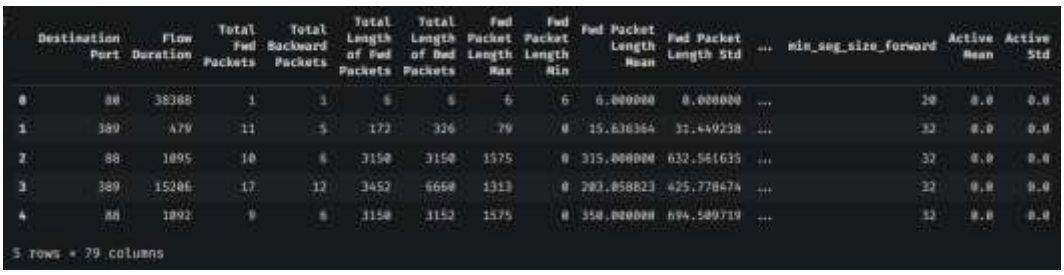


| Label | |
|---|---|
| BENIGN | 682092 |
| DoS Hulk | 69066 |
| PortScan | 47841 |
| DDoS | 38346 |
| DoS GoldenEye | 8841 |
| FTP-Patator | 2359 |
| SSH-Patator | 1786 |
| DoS slowloris | 1764 |
| DoS Slowhttptest | 1649 |
| Bot | 580 |
| Web Attack  Brute Force | 465 |
| Web Attack  XSS | 203 |
| Infiltration | 21 |
| Web Attack  Sql Injection | 8 |
| Heartbleed | 7 |

Name: count, dtype: int64

**Gambar1 machine learning maupun deep learning.**

Pada penelitian ini digunakan versi *dataset* yang telah diproses menjadi fitur-fitur numerik sehingga sesuai untuk proses *preprocessing*, normalisasi, transformasi menjadi representasi teks, pembentukan *embedding* menggunakan *Sentence-BERT*, serta seleksi fitur berbasis NLP. Setelah proses pembersihan data, atribut yang digunakan terdiri atas puluhan fitur numerik dan satu atribut label yang menunjukkan kategori lalu lintas jaringan, yaitu *Benign* maupun berbagai jenis serangan. Selanjutnya, seluruh fitur numerik dinormalisasi menggunakan metode Min-Max *Scaling* sebelum dikonversi menjadi kalimat menggunakan *template* bahasa alami yang konsisten.



| | Destination Port | Flow Duration | Total Fwd Packets | Total Backward Packets | Total Length of Fwd Packets | Total Length of Bwd Packets | Fwd Packet Length Max | Fwd Packet Length Min | Fwd Packet Length Mean | Fwd Packet Length Std | ... | min_seg_size_forward | Active Mean | Active Std |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 80 | 38388 | 1 | 1 | 6 | 6 | 6 | 6 | 6.000000 | 0.000000 | ... | 20 | 0.0 | 0.0 |
| 1 | 389 | 479 | 11 | 5 | 172 | 326 | 70 | 0 | 15.636364 | 31.449258 | ... | 32 | 0.0 | 0.0 |
| 2 | 88 | 1095 | 10 | 6 | 3150 | 3150 | 1575 | 0 | 315.000000 | 632.561635 | ... | 32 | 0.0 | 0.0 |
| 3 | 389 | 15206 | 17 | 12 | 3452 | 6660 | 1313 | 0 | 203.058823 | 425.778474 | ... | 32 | 0.0 | 0.0 |
| 4 | 88 | 1092 | 9 | 6 | 3150 | 3152 | 1575 | 0 | 350.000000 | 694.509719 | ... | 32 | 0.0 | 0.0 |

5 rows × 79 columns

**Gambar 2 fitur-fitur numerik**

Melalui proses ini, seluruh fitur memiliki skala yang sebanding tanpa mengubah hubungan relatif antardata. Selain meningkatkan konsistensi representasi data, normalisasi juga membantu proses pembentukan *embedding* karena model bahasa tidak dipengaruhi oleh perbedaan besaran nilai numerik yang ekstrem. Setelah seluruh fitur dinormalisasi, data tidak langsung digunakan sebagai masukan model klasifikasi. Berbeda dengan pendekatan konvensional yang mempertahankan bentuk tabel numerik, penelitian ini mengubah setiap *record* menjadi representasi bahasa alami (*natural language representation*). Transformasi tersebut dilakukan menggunakan *template* kalimat yang konsisten sehingga setiap atribut jaringan dapat diperlakukan sebagai informasi tekstual yang memiliki konteks semantik.

| | Destination Port | Flow Duration | Total Fwd Packets | Total Backward Packets | Total Length of Fwd Packets | Total Length of Bwd Packets | Fwd Packet Length Max | Fwd Packet Length Min | Fwd Packet Length Mean | Fwd Packet Length Std | ... | act_data_pkt_fwd | min_seg_size_ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.001222 | 0.000319 | 0.000000 | 0.000004 | 0.000005 | 9.589178e-05 | 0.000242 | 0.002986 | 0.001293 | 0.000000 | ... | 0.000000 | 0 |
| 1 | 0.005940 | 0.000004 | 0.000049 | 0.000018 | 0.000141 | 5.199362e-07 | 0.003183 | 0.000000 | 0.003369 | 0.004092 | ... | 0.000020 | 0 |
| 2 | 0.001344 | 0.000009 | 0.000044 | 0.000023 | 0.002573 | 5.023023e-06 | 0.063457 | 0.000000 | 0.067077 | 0.098389 | ... | 0.000015 | 0 |
| 3 | 0.003940 | 0.000127 | 0.000070 | 0.000044 | 0.002820 | 1.002201e-05 | 0.032981 | 0.000000 | 0.041750 | 0.066226 | ... | 0.000051 | 0 |
| 4 | 0.001344 | 0.000009 | 0.000039 | 0.000023 | 0.002573 | 5.027113e-06 | 0.063457 | 0.000000 | 0.075419 | 0.100814 | ... | 0.000010 | 0 |

5 rows × 78 columns

**Gambar 3 konteks semantik**

**Seleksi Fitur** Setelah setiap *record* pada *dataset*

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
August 2026, IX (4): 6124 – 6132 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

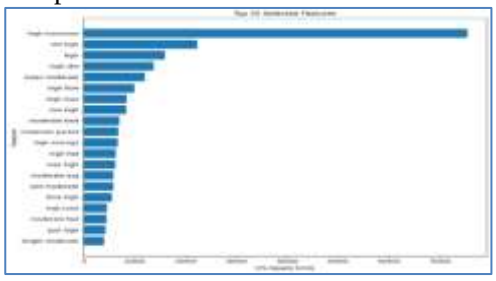
CICIDS2017 ditransformasikan *menjadi* representasi bahasa alami menggunakan *template* kalimat yang konsisten, tahapan lanjutan adalah melakukan seleksi fitur berbasis NLP. Berdasarkan hasil dari transformasi data, diperoleh *dataset* dengan ukuran 691406 baris dan 2 kolom. Jumlah baris tersebut menunjukkan banyaknya *network flow* atau *record* yang berhasil diproses. Sementara dua kolom yang tersedia terdiri atas kolom teks, yang berisi representasi dari setiap *flow* jaringan hasil transformasi menggunakan *template* kalimat NLP, dan kolom tabel yang menunjukkan kategori lalu lintas jaringan, yaitu *Benign* atau salah satu jenis serangan.

Dengan jumlah 691.406 dokumen, *dataset* memiliki ukuran yang cukup besar untuk membangun representasi fitur yang kaya melalui proses pembobotan TF-IDF, seleksi fitur menggunakan Chi-Square, maupun pembentukan *embedding* menggunakan model berbasis *Transformer*. Selain itu, ukuran *dataset* yang besar diharapkan mampu meningkatkan kemampuan model dalam mempelajari pola lalu lintas normal dan berbagai jenis serangan sehingga menghasilkan sistem deteksi intrusi yang lebih akurat dan memiliki kemampuan generalisasi yang baik.

Dalam penelitian ini, proses seleksi fitur dilakukan melalui dua tahap, yaitu pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan pemilihan fitur menggunakan metode *Chi-Square* (Chi2). Penggunaan kombinasi dari kedua metode ini bertujuan untuk memperoleh fitur tekstual yang paling relevan untuk membedakan lalu lintas normal dan serangan.

**Gambar 4 hasil ekstraksi fitur**

Dengan demikian, matriks TF-IDF berukuran (691406, 302) menjadi dasar dalam proses seleksi fitur dan klasifikasi. Representasi ini mengubah data *tabular* menjadi ruang vektor berdimensi 302 yang tetap mempertahankan informasi penting dari setiap *network flow*, sehingga dapat dimanfaatkan secara efektif oleh algoritma *machine learning* dalam mendeteksi intrusi pada jaringan komputer.



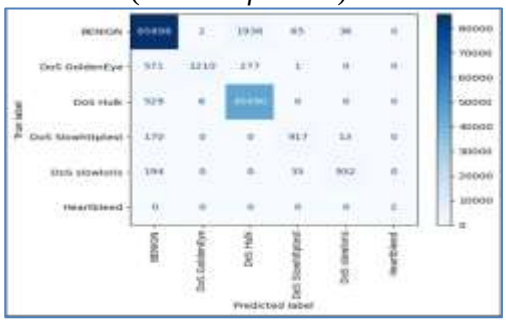
**Gambar 5 algoritma *machine learning***

Gambar tersebut menunjukkan 20 fitur (term) terbaik hasil seleksi menggunakan metode *Chi-Square* ($Chi^2$) berdasarkan bobot TF-IDF. Sumbu horizontal menunjukkan nilai *Chi-Square Score*, sedangkan sumbu vertikal menunjukkan term yang terpilih. Semakin tinggi nilai Chi-Square, semakin besar kontribusi term tersebut dalam membedakan lalu lintas normal dan serangan. Terlihat bahwa term "*high maximum*" memiliki skor tertinggi, sehingga menjadi fitur yang paling berpengaruh dalam proses klasifikasi. Hasil ini menunjukkan bahwa metode Chi-Square berhasil mengidentifikasi term-term yang paling relevan, sehingga hanya fitur yang memiliki kemampuan diskriminatif tinggi yang digunakan pada tahap klasifikasi untuk meningkatkan

efisiensi dan mempertahankan kinerja model deteksi intrusi.

**Analisis Confusion Matrix**

*Confusion matrix* digunakan untuk mengevaluasi kinerja model dengan membandingkan kelas sebenarnya (*true label*) dan kelas hasil prediksi (*predicted label*). Nilai yang berada pada diagonal utama menunjukkan jumlah data yang berhasil diklasifikasikan dengan benar (*True Positive*), sedangkan nilai di luar diagonal menunjukkan kesalahan klasifikasi (*misclassification*).




**Gambar 6 Analisis Confusion Matrix**

Untuk kelas *DoS Hulk*, model menunjukkan performa yang sangat baik dengan 45.490 data yang berhasil dikenali secara benar. Kesalahan klasifikasi relatif kecil, yaitu 529 data diprediksi sebagai BENIGN dan 6 data sebagai *DoS GoldenEye*. Tingginya jumlah prediksi benar menunjukkan bahwa fitur-fitur hasil seleksi mampu merepresentasikan karakteristik serangan *DoS Hulk* dengan baik, sehingga mudah dikenali oleh algoritma *Random Forest*.

Pada kelas *DoS Slowhttptest*, model berhasil mengidentifikasi 917 data dengan benar. Kesalahan yang terjadi meliputi 170 data yang diprediksi sebagai BENIGN dan 13 data sebagai DoS slowloris. Demikian pula, pada kelas DoS Slowloris, sebanyak 932 data berhasil diklasifikasikan dengan benar, sedangkan 194 data diprediksi sebagai BENIGN dan 33 data sebagai DoS Slowhttptest. Kesalahan klasifikasi antara kedua kelas tersebut cukup wajar karena keduanya termasuk jenis serangan DoS berbasis koneksi lambat (low-rate DoS attack) yang memiliki karakteristik lalu lintas jaringan yang hampir serupa.

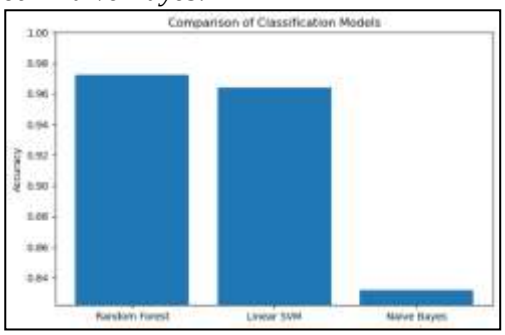
Kelas Heartbleed menunjukkan hasil yang sangat baik, di mana kedua data berhasil diklasifikasikan dengan benar tanpa terjadi kesalahan prediksi. Namun demikian, jumlah sampel Heartbleed pada dataset sangat sedikit sehingga hasil ini belum cukup untuk menyimpulkan bahwa model memiliki kemampuan yang sangat baik dalam mendeteksi serangan tersebut. Ketidakseimbangan jumlah data (class imbalance) perlu menjadi perhatian karena dapat memengaruhi kemampuan model dalam mengenali kelas yang jumlah datanya sangat kecil.

Secara keseluruhan, *confusion matrix* menunjukkan bahwa nilai terbesar berada pada diagonal utama, yang menandakan sebagian besar data berhasil diklasifikasikan sesuai dengan kelas sebenarnya. Hal ini mengindikasikan bahwa model *Random Forest* memiliki kemampuan klasifikasi yang tinggi setelah menggunakan fitur hasil seleksi berbasis TF-IDF dan *Chi-Square*. Meskipun masih terdapat beberapa kesalahan klasifikasi, terutama antara kelas *Benign*, *DoS GoldenEye*, dan *DoS Hulk*, jumlah kesalahan tersebut relatif kecil dibandingkan dengan jumlah prediksi yang benar.

Hasil ini membuktikan bahwa proses transformasi data numerik menjadi representasi teks menggunakan NLP, diikuti dengan pembobotan TF-IDF dan seleksi fitur menggunakan *Chi-Square*, mampu menghasilkan fitur-fitur yang informatif untuk proses klasifikasi. Fitur-fitur tersebut membantu algoritma *Random Forest* dalam membedakan karakteristik berbagai jenis lalu lintas jaringan sehingga menghasilkan performa deteksi yang tinggi pada *dataset* CICIDS2017. Dengan demikian, pendekatan yang diusulkan efektif untuk meningkatkan kemampuan NIDS dalam mendeteksi serangan siber dengan tingkat akurasi yang tinggi sekaligus mengurangi kompleksitas data melalui proses seleksi fitur.

**Evaluasi Performa**

Tahapan terakhir yang dilakukan adalah evaluasi terhadap performa model klasifikasi. Pada tahapan ini digunakan tiga algoritma klasifikasi, yaitu *Random Forest, Linear Support Vector Machine*, dan *Naive Bayes*.




**Gambar 7 Evaluasi Performa**

Terlihat bahwa *Random Forest* memperoleh akurasi tertinggi, yaitu sekitar 97,3%, diikuti oleh Linear SVM dengan akurasi sekitar 96,4%. Sementara itu, *Naive Bayes* hanya mencapai akurasi sekitar 83,2%, sehingga memiliki kinerja yang jauh lebih rendah dibandingkan kedua algoritma lainnya. Hasil ini menunjukkan bahwa *Random Forest* merupakan model yang paling efektif dalam mengklasifikasikan lalu lintas jaringan pada *dataset* CICIDS2017 karena mampu memanfaatkan fitur hasil seleksi dengan lebih baik. Sebaliknya, performa *Naive Bayes* yang lebih rendah mengindikasikan bahwa asumsi independensi antarfitur pada algoritma tersebut kurang sesuai dengan karakteristik data hasil representasi NLP. Oleh karena itu, *Random Forest* dipilih sebagai model terbaik karena memberikan tingkat akurasi tertinggi dalam mendeteksi baik lalu lintas normal maupun berbagai jenis serangan jaringan.

## SIMPULAN

Berdasarkan penelitian yang telah dilakukan, usulan pendekatan BERT-NIDS yang mengintegrasikan representasi semantik berbasis NLP, pembobotan TF-IDF, dan seleksi fitur *Chi-Square* terbukti efektif dalam meningkatkan efisiensi dan performa NIDS pada *dataset* CICIDS2017. Transformasi data numerik menjadi representasi bahasa alami berhasil mereduksi kompleksitas data secara signifikan sekaligus mempertahankan informasi penting untuk mendeteksi berbagai jenis serangan siber. Melalui pengujian pada beberapa algoritma klasifikasi, *Random Forest* terbukti menjadi model terbaik dengan pencapaian akurasi tertinggi sekitar 97,3%, mengungguli Linear SVM dan *Naive Bayes*, sehingga pendekatan ini layak dipertimbangkan sebagai solusi yang andal dan akurat untuk sistem deteksi intrusi jaringan modern.

## DAFTAR PUSTAKA

Yusda, R. A., Sahren, S., Fitri Larasti Sibuea, M., Meutia Arifin, N., & Aditya, B. (2025). Seleksi Fitur Menggunakan Mutual Information Untuk Deteksi Intrusi. *Journal of Science and Social Research*, *8*(3), 3482–3490. http://jurnal.goretanpena.com/index.php/JSSR

Arreche, O., Guntur, T., & Abdallah, M. (2024). XAI-based Feature Selection for Improved Network Intrusion Detection Systems. *ArXiv*. http://arxiv.org/abs/2410.10050

Cybersecurity Ventures. (2023). *Cybersecurity Almanac*.

Ma, H., Zhang, W., Zhang, D., & Chen, B. (2025). An IoT intrusion detection framework based on feature selection and large language models fine-tuning. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-08905-3

Maasaoui, Z., Battou, A., Merzouki, M., & Lbath, A. (2024, July). *Anomaly Based Intrusion Detection using Large Language Models*.

Mamikonian, V., & Mamamniashvili, D. (2025). A Lightweight and Optimized BERT-based Intrusion Detection System for Resource-Constrained Network and Logistics

Environments. *GEORGIAN SCIENTISTS*. https://doi.org/10.52340/gs.2025.07.04.49

Ngo, V.-D., Vuong, T.-C., Van Luong, T., & Tran, H. (2023). Machine Learning-Based Intrusion Detection: Feature Selection versus Feature Extraction. *ArXiv*. http://arxiv.org/abs/2307.01570

Ponemon Institute. (2024). *Third Annual Ponemon Institute Report_ Nearly Seven in 10 Healthcare Organizations Experienced Disruption to Patient Care Due to Cyber Attacks _ Proofpoint US*.

Sahren, S., & Prijuna Lubis, A. (2024). Intrusion Detection System Berbasis Deep Learning Untuk Peningkatan Mitigasi SQL Injection dan Syn Flood Attack. *Journal of Science and Social Research*, *7*(4), 1866–1874. http://jurnal.goretanpena.com/index.php/JSSR

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSP 2018 - Proceedings of the 4th International Conference on Information Systems Security and Privacy*, *2018-January*, 108–116. https://doi.org/10.5220/0006639801080116

Sworna, Z. T., Mousavi, Z., & Babar, M. A. (2023). NLP methods in host-based intrusion detection systems: A systematic review and future directions. In *Journal of Network and Computer Applications* (Vol. 220). Academic Press. https://doi.org/10.1016/j.jnca.2023.103761

Talukder, M. A., Hasan, K. F., Islam, M. M., Uddin, M. A., Akhter, A., Yousuf, M. A., Alharbi, F., & Moni, M. A. (2023). A dependable hybrid machine learning model for network intrusion detection. *Journal of Information Security and Applications*, *72*, 103405. https://doi.org/10.1016/J.JISA.2022.103405

Umer, M., Tahir, M., Sardaraz, M., Sharif, M., Elmannai, H., & Algarni, A. D. (2025). Network intrusion detection model using wrapper based feature selection and multi head attention transformers. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-11348-5

Vubangsi, M., Mangai, T. R., Olukayode, A., Mubarak, A. S., & Al-Turjman, F. (2024). BERT-IDS: an intrusion detection system based on bidirectional encoder representations from transformers. *Computational Intelligence and Blockchain in Complex Systems: System Security and Interdisciplinary Applications*, 147–155. https://doi.org/10.1016/B978-0-443-13268-1.00021-2

Westphal, C., Hailes, S., & Musolesi, M. (2024). Feature Selection for Network Intrusion Detection. *ArXiv*. http://arxiv.org/abs/2411.11603

Yang, Y., & Peng, X. (2025). BERT-based network for intrusion detection system. *Eurasip Journal on Information Security*, *2025*(1). https://doi.org/10.1186/s13635-025-00191-w

Yin, Y., Jang-Jaccard, J., Xu, W., Singh, A., Zhu, J., Sabrina, F., & Kwak, J. (2023). IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. *Journal of Big Data*, *10*(1). https://doi.org/10.1186/s40537-023